A Bayesian General Linear Modeling Approach to Cortical Surface fMRI Data Analysis

A Bayesian General Linear Modeling Approach to Cortical Surface fMRI Data Analysis

Amanda F. Mejia[a*], Yu Ryan Yue[b*], David Bolin[c], Finn Lindgren[d] and Martin A. Lindquist[e]

[a] *Indiana University, Bloomington, IN 47405*
[b] *Baruch College, The City University of New York, New York, NY 10010*
[c] *University of Gothenburg, Gothenburg, Sweden*
[d] *The University of Edinburgh, Edinburgh, UK*
[e] *Johns Hopkins University, Baltimore, MD 21205*

[*] A. Mejia and Y. Yue contributed equally to this work.

Abstract

Cortical surface fMRI (cs-fMRI) has recently grown in popularity versus traditional volumetric fMRI. In addition to offering better whole-brain visualization, dimension reduction, removal of extraneous tissue types, and improved alignment of cortical areas across subjects, it is also more compatible with common assumptions of Bayesian spatial models. However, as no spatial Bayesian model has been proposed for cs-fMRI data, most analyses continue to employ the classical general linear model (GLM), a "massive univariate" approach. Here, we propose a spatial Bayesian GLM for cs-fMRI, which employs a class of sophisticated spatial processes to model latent activation fields. We make several advances compared with existing spatial Bayesian models for volumetric fMRI. First, we use integrated nested Laplacian approximations (INLA), a highly accurate and efficient Bayesian computation technique, rather than variational Bayes (VB). To identify regions of activation, we utilize an excursions set method based on the joint posterior distribution of the latent fields, rather than the marginal distribution at each location. Finally, we propose the first multi-subject spatial Bayesian modeling approach, which addresses a major gap in the existing literature. The methods are very computationally advantageous and are validated through simulation studies and two task fMRI studies from the Human Connectome Project.

Keywords: spatial statistics; Bayesian smoothing; integrated nested Laplace approximation; stochastic partial differential equation; brain imaging

1 INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a popular noninvasive neuroimaging technique commonly used to localize regions of the brain activated by a task or stimulus (Lindquist 2008, Poldrack et al. 2011). Traditional *volumetric* fMRI data consists of a time series of 3-dimensional brain volumes, each composed of thousands of equally sized volumetric elements (voxels). While neuronal activity is known to occur in gray matter, volumetric fMRI includes several other tissue classes, including white matter and cerebral spinal fluid. Cortical surface fMRI (cs-fMRI) is an alternative representation in which the cortical gray matter is represented as a 2-dimensional manifold surface (Fischl 2012, Glasser et al. 2013). Cs-fMRI offers several advantages, including whole-brain visualization, dimension reduction, removal of extraneous tissue types, improved alignment across subjects, and—importantly for analysis purposes—greater neurobiological significance of distances (Figure 1). That is, nearby locations along the cortex tend to exhibit similar patterns of neuronal activity, while in volumetric fMRI nearby locations may be neurobiologically quite dissimilar (Figure 2). Cs-fMRI is therefore highly advantageous for spatial models, which pool information across neighbors.

The process of transforming volumetric to cs-fMRI is illustrated in Figure 1. First, a high-resolution structural image is used to identify the cortical gray matter ribbon (Dale et al. 1999). Second, a mesh is applied to the white matter surface, the internal boundary of the cortical gray matter, to form a 2-dimensional manifold within each hemisphere, which is then geometrically smoothed. Third, the surface is inflated to a sphere while minimizing distance distortions, and subjects are registered to a standard template space by aligning anatomical folding patterns (Fischl et al. 1999). Finally, for each fMRI volume the same volume-to-surface transformation is applied to the cortical gray matter ribbon to obtain a cs-fMRI time series. The resulting data for each volume is a triangular mesh consisting of approximately 30,000 vertices within each hemisphere. Once registration has been performed, this mesh can be de-inflated to various levels to facilitate visualization or modeling.

The traditional task analysis method for both volumetric and cs-fMRI data is the classical *general linear model* (GLM), also known as the "massive univariate" approach, in which a linear regression model relating the observed fMRI data to the expected BOLD response to each task is fit separately at each voxel or vertex (Friston et al. 1994). Although it has long been

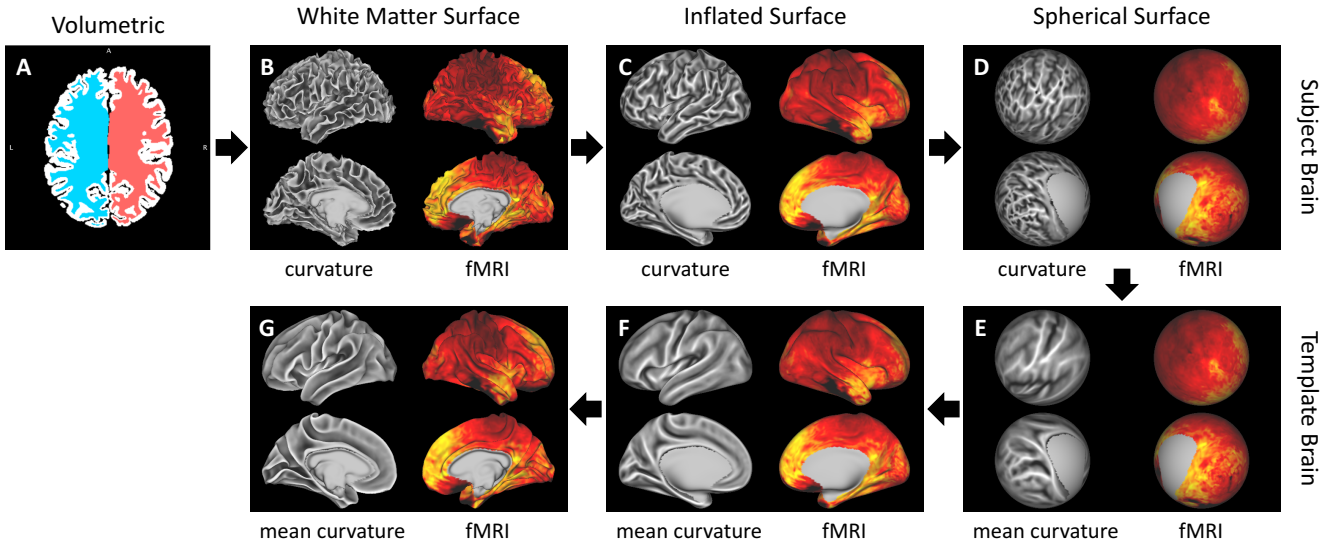

Figure 1: Illustration of transformation from volumetric to surface representation. **(A)** In a single slice of volumetric space, the white matter in the left (blue) and right hemisphere (red) and the cortical gray matter (white). **(B)-(G)** In surface space, the four-way view showing the *lateral* or exterior (top) and *medial* or interior (bottom) views of both hemispheres. Here we display curvature values on the left hemisphere and the BOLD response from a single fMRI volume on the right hemisphere; in both cases, lighter colors represent larger values. **(B)** The white matter surface of a subject after a mesh has been applied to the white matter boundary and smoothed. **(C)** and **(D)** Two different levels of inflation of the white matter surface. **(D)** to **(E)** Subject brains are aligned to the template brain by aligning cortical folding patterns, indicated by curvature, on the spherical surface. **(F)** and **(G)** The template brain at lower levels of inflation, which are often preferable for display purposes.

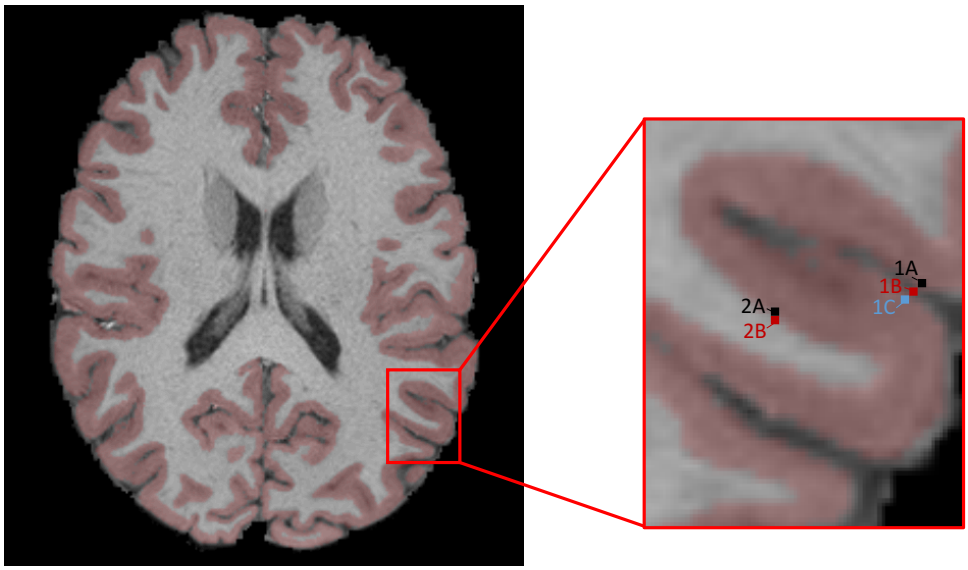

Figure 2: Distances in volumetric space. For one subject, an axial slice of the T_1-weighted image is displayed, with the cortical gray matter overlaid in red. Locations 1A, 1B and 1C are close in terms of Euclidean distance in volumetric space, but are neurologically dissimilar, as location 1A lies on one sulcal bank, location 1B lies in the cerebrospinal fluid between sulcal banks, and location 1C lies on an opposite sulcal bank. Therefore, locations 1A and 1C may exhibit distinct task activation patterns, while location 1B would not be expected to exhibit any task-related activation. Similarly, locations 2A and 2B are neighboring in volumetric space, but location 2A lies in the cortical gray matter while location 2B lies in the white matter and therefore would not be expected to exhibit task-related activation. This illustrates the limitations of the volumetric representation for task fMRI analysis: the classical GLM model typically employs smoothing with a Gaussian kernel throughout the volume, which would have the result of mixing the distinct signals from locations 1A, 1B and 1C (and those from locations 2A and 2B), while a Bayesian approach assuming a stationary prior on the latent task fields would incorrectly assume the latent signal at locations 1A, 1B and 1C (and locations 2A and 2B) to be highly correlated. By contrast, in the cortical surface representation, locations 1B and 2B would be excluded from analysis, as they do not lie within the cortical gray matter, and the latent signals at locations 1A and 1C would be assumed to have low dependence, due to the greater geodesic distance along the cortical surface between them.

recognized that individual voxels and vertices do not live in a vacuum, but rather present very similar activation patterns as tens or hundreds of other locations within a surrounding region, the classical GLM was initially proposed as a feasible way to analyze data which otherwise would have presented insurmountable computational challenges. We now describe the steps involved in the classical GLM and describe their limitations before discussing the challenges to successful spatial Bayesian modeling and our proposed solutions to these challenges.

1.1 The classical GLM approach

Prior to model fitting in the classical GLM, the fMRI data is typically smoothed using a fixed-width Gaussian kernel in order to increase the signal-to-noise ratio (SNR) of the data. The coefficients of the model are estimates of the task-related activation at each location, the significance of which is tested using a t or F statistic. The corresponding p-values are then plotted at each location to form a *statistical parametric map* (SPM). To identify the areas of true activation, the null hypothesis of no activation at each location is tested by thresholding the SPM at significance level α (Worsley and Friston 1995), chosen to control the family-wise error rate (FWER) or false discovery rate (FDR) at some predetermined level (Genovese et al. 2002, Lindquist and Mejia 2015). To account for correlations between tests performed at neighboring locations, popular solutions include parametric methods such as *random field theory* (RFT) (Adler 1981), Monte Carlo simulation (Forman et al. 1995), and nonparametric methods such as permutation tests (Nichols and Holmes 2002). To avoid identifying very small regions of spurious activation, cluster-based methods first threshold the SPM at a fixed level determined by the researcher (e.g. $p = 0.01$ or $p = 0.001$) then use permutation tests, random field theory, or ad-hoc methods to determine significant clusters (Poline and Mazoyer 1993). Threshold-free cluster enhancement (Smith and Nichols 2009) approaches circumvent the dependence on an arbitrary cluster-forming threshold.

While the effort to properly correct for multiple comparisons in the classical GLM is a necessary one, most of the traditional correction methods have been shown to suffer from various pitfalls. In high-dimensional settings, methods that control the FWER or FDR have been shown to suffer from a lack of power to detect true effects (Ishwaran and Rao 2003, Marchini and Presanis 2004). Parametric methods such as RFT have been found to be inaccurate due to de-

partures of the data from the parametric assumptions (Nichols and Hayasaka 2003, Wager et al. 2009, Eklund et al. 2012; 2016). Further, as voxels (and vertices) lack biological meaning as a unit of measure, controlling the voxel-wise FWER is rather arbitrary and is sensitive to voxel size, which is gradually shrinking due to technological advances. While cluster-based methods avoid this limitation, they have been found to be sensitive to the choice of initial threshold (Woo et al. 2014, Eklund et al. 2016). Finally, since the corrected significance threshold becomes more conservative as the number of tests increases, inference is sensitive to the size of the search volume.

These issues are symptomatic of some of the fundamental limitations of the classical GLM. First, while it is well-known that the activation amplitude of one voxel depends on its neighbors, the classical GLM does not account for such dependence. Second, while spatial smoothing of the fMRI data prior to model fitting can increase SNR and help satisfy the assumptions of RFT, when applied to volumetric fMRI data it may also combine signal from different tissue types and across discontiguous regions of the cortex (see Figure 2), contaminating the signal of interest and leading to inaccurate identification of truly active regions. Smoothing of cs-fMRI data is less problematic but also tends to blur boundaries between active and non-active areas, and a common degree of smoothing may not be optimal for each latent activation field. Data smoothing also results in smoothed noise, thus increasing dependence between tests and complicating the problem of correcting for multiple comparisons.

1.2 Existing spatial Bayesian approaches

Spatial Bayesian models have long held promise as a solution to these issues (Friston and Penny 2003). In a Bayesian GLM, specific prior distributions are assumed for the latent task activation fields and other unknown parameters in the model, and they, together with the likelihood, form a Bayesian hierarchical model. For each location, the posterior probability that the corresponding amplitude is greater than some biologically meaningful *activation threshold* (often a percentage of global mean signal) is calculated, and active locations are identified by thresholding the resulting *posterior probability map* (PPM) at a certain level (e.g., 0.95).

However, these models have not seen wide adoption for task fMRI analysis. One of the main issues is that while all existing spatial Bayesian models for fMRI have been designed for

the regular lattice structure of volumetric fMRI, the complex spatial dependence structure of volumetric fMRI described above is inconsistent with the modeling assumptions (see Figure 2). This presents a fundamental issue that cannot be easily resolved without moving away from Euclidean distances within the brain volume. Here, we propose a spatial Bayesian modeling approach for cs-fMRI data, which utilizes geodesic distances along the cortical surface. This is, to the best of our knowledge, the first spatial Bayesian modeling approach for cs-fMRI. We first describe several serious challenges to successful spatial Bayesian modeling of fMRI data, which our proposed approach addresses.

First, volumetric fMRI data consists of 100,000-200,000 voxels each measured at hundreds of time points, presenting serious computational challenges. To surmount these challenges, it is typical to use variational Bayesian (VB) techniques for the Bayesian computation (Penny et al. 2005, Sidén et al. 2017), which are known to severely underestimate posterior variance and may provide poor estimates of the posterior mode (Wang and Titterington 2005, Bishop 2006, Rue et al. 2009, Sidén et al. 2017). Furthermore, most existing spatial Bayesian models are designed to be fit within each 2-dimensional slice of the brain volume separately, introducing discontinuities in the resulting activation estimates and regions (Sidén et al. 2017). Cs-fMRI can somewhat alleviate this burden, as it contains approximately 30,000 vertices per hemisphere and can be resampled to a lower resolution without great loss of information, unlike volumetric fMRI.

Second, in order to identify activations, for computational feasibility the marginal, rather than joint, posterior distribution at each location is typically used. This introduces a multiple comparisons problem that must be corrected, resulting in potential loss of power. Finally, again due to computational limitations, all existing spatial Bayesian models for volumetric fMRI data have been designed for single-subject analysis. As group-level inference is often a primary goal in task fMRI studies, this represents a major limitation of existing methods (Sidén et al. 2017).

1.3 Proposed spatial Bayesian approach for cs-fMRI

We now briefly describe our proposed approach for spatial Bayesian modeling of cs-fMRI and how each of the issues described above is addressed. The main challenges are (1) selecting a prior on latent task activation fields that is appropriate for cs-fMRI data, (2) performing

the Bayesian computation accurately yet efficiently, (3) identifying activations using the joint posterior distribution of each latent field, and (4) performing group-level inference.

To account for the spatial dependence in activation levels, the model coefficients are assumed to follow a spatial process prior, which must have a sparse inverse covariance structure. Priors advocated for volumetric fMRI are typically designed for data on a regular lattice and may not be applicable to cs-fMRI data. Here, we employ a class of GMRF priors introduced by Lindgren et al. (2011) termed stochastic partial differential equations (SPDE) priors, which possess several advantages for analyzing fMRI data. First, unlike regular GMRF priors, SPDE priors are explicit mappings of continuous Matérn Gaussian fields, which have been extensively used in statistical modeling of spatial data (e.g., Guttorp and Gneiting 2006). Therefore, they combine the computational advantages of a GMRF with the flexibility of the Matérn covariance structure and the interpretability of the Matérn parametrization. Second, SPDE priors are constructed on a flexible triangular mesh, the structure of cs-fMRI data. The triangular mesh structure allows for appropriate smoothing along boundaries by emphasizing neighbors within the boundary and de-emphasizing neighbors across the boundary, thus avoiding boundary blurring. Finally, SPDE priors are consistent under re-triangulations of the surface, which is not true in general of GMRF models that are not defined as discretisations of continuous models (Lindgren et al. 2011).

For the Bayesian computation, we employ a recently developed Bayesian inference tool based on integrated nested Laplace approximations (INLA) (Rue et al. 2009). The INLA method can directly compute very good approximations to the posterior distributions and is able to handle large data sets by taking advantage of the sparsity of GMRFs. It tends to be faster than MCMC (Rue et al. 2016) since it does not require sampling, and can be easily implemented using the `R-INLA` package (Martins et al. 2013). To identify areas of activation based on the joint, rather than the marginal, posterior probabilities, we use the excursions set method introduced by Bolin and Lindgren (2015), which uses INLA to estimate a joint posterior probability map (PPM) that can be thresholded to identify a set of locations that are activated at a given probability level. This avoids the need for multiple comparisons correction and fully leverages spatial dependencies at each step in the model.

Finally, we propose the first multi-subject spatial Bayesian modeling approach for fMRI

analysis. Due to the computational advantages of the proposed methods, it is in some cases feasible to directly perform multi-subject analysis through a single group model. However, in many cases fitting such a model may be infeasible. Therefore, we propose a novel approach for combining the results of individual subject-level models. Using this approach, we can estimate the posterior distribution of each group-level latent activation field to obtain estimates and areas of activation.

The remainder of this paper is organized as follows. The Bayesian GLM method is introduced in Section 2, where the SPDE priors, INLA algorithm and joint PPM approach are presented. The method is then extended to multi-subject analysis in Section 3. We assess the accuracy of the proposed Bayesian methods in a simulation study described in Section 4, followed by an application to two task fMRI studies in Section 5. We conclude with a discussion in Section 6.

2 SINGLE SUBJECT BAYESIAN GLM

Let T be the number of time points in the fMRI timeseries and let N be the number of vertices in each hemisphere of the brain. For a subject and hemisphere, we have the following model:

$$\boldsymbol{y} = \sum_{k=0}^{K} \boldsymbol{X}_k \boldsymbol{\beta}_k + \sum_{j=1}^{J} \boldsymbol{Z}_j \boldsymbol{b}_j + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \boldsymbol{V}). \tag{1}$$

Here $\boldsymbol{y}$ is an $TN \times 1$ vector containing the fMRI time series of all vertices, and the $\boldsymbol{X}_k$ and $\boldsymbol{Z}_j$ are $TN \times N$ design matrices for the activation amplitudes $\boldsymbol{\beta}_k$ (including baseline $\boldsymbol{\beta}_0$) and nuisance signals $\boldsymbol{b}_j$, respectively. The matrix $\boldsymbol{V} = \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}(\xi, \boldsymbol{\phi})$, where $\boldsymbol{\Sigma}(\xi, \boldsymbol{\phi})$ is a $T \times T$ covariance matrix for an AR(p) process with *marginal* precision ξ and partial autocorrelation functions $\boldsymbol{\phi} = (\phi_1, , \ldots, \phi_p)'$ assumed for each time series, and $\otimes$ denotes the Kronecker product. For fully Bayesian inference, prior distributions are assumed on the unknown parameters in model (1). For nuisance parameters in $\boldsymbol{b}_j$, we take independent and diffuse Gaussian priors, that is $\boldsymbol{b}_j \sim N(\boldsymbol{0}, \delta \boldsymbol{I})$ where δ is a fixed large number. For ξ and $\boldsymbol{\phi}$, we first reparameterize them as

$$\theta_1 = \log(\xi), \quad \theta_r = \log\left(\frac{1 + \phi_{r-1}}{1 - \phi_{r-1}}\right),$$

for $r = 2, \ldots, p+1$. We then let θ_1 follow a diffuse log gamma prior and $(\theta_2, \ldots, \theta_{p+1})$ a multivariate normal prior. The spatial priors for $\boldsymbol{\beta}_k$ are described below.

2.1 SPDE Spatial Priors

To account for spatial homogeneity, we need to take a spatial prior on each $\boldsymbol{\beta}_k$ for $k = 0, \ldots, K$. A good candidate is the class of Matérn Gaussian fields that has been extensively used in spatial statistics due to its flexible covariance function between locations. We say $\beta(\boldsymbol{u})$ is a Matérn Gaussian process if the covariance between $\boldsymbol{u}$ and $\boldsymbol{v}$ $(\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d)$ is given by

$$\text{Cov}(\boldsymbol{u}, \boldsymbol{v}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa\|\boldsymbol{u} - \boldsymbol{v}\|)^\nu K_\nu(\kappa\|\boldsymbol{u} - \boldsymbol{v}\|),$$

where $K_\nu(\cdot)$ is the modified Bessel function of the second kind with order $\nu > 0$, where ν controls the smoothness; $\|\cdot\|$ is the Euclidean norm; $\Gamma(\cdot)$ is the gamma function; $\kappa > 0$ is the spatial scale; and $\sigma^2 > 0$ is the variance. However, a Matérn spatial process is not computationally feasible for large data sets because its covariance matrix is completely dense and therefore difficult to invert. Lindgren et al. (2011) addressed this issue by deriving an explicit GMRF representation for Matérn Gaussian fields through solving the following SPDE

$$\left(\kappa^2 - \Delta\right)^{\alpha/2} (\tau\beta(\boldsymbol{u})) = \mathcal{W}(\boldsymbol{u}), \quad \boldsymbol{u} \in \mathbb{R}^d \tag{2}$$

where $\Delta = \sum_{i=1}^d \partial^2/\partial u_i^2$ is the Laplacian operator, α affects the smoothness, and τ affects the variance. $\mathcal{W}(\boldsymbol{u})$ is the Gaussian white noise process. The stationary solution β to this SPDE is a Matérn Gaussian field, and the link to the smoothness ν and variance σ^2 is $\nu = \alpha - d/2$ and $\sigma^2 = \Gamma(\nu)\left[\Gamma(\alpha)(4\pi)^{d/2}\kappa^{2\nu}\tau^2\right]^{-1}$. As spectral theory shows that an integer α must be chosen to obtain a Markov field, we let $\alpha = 2$, resulting in $\nu = 1$ for a two-dimensional field.

To obtain a Markov structure, we approximate $\beta(\boldsymbol{u})$ using the basis expansion

$$\beta(\boldsymbol{u}) \approx \sum_{i=1}^n \psi_i(\boldsymbol{u}) w_i. \tag{3}$$

Here ψ_i is a piecewise linear basis function taking a value of 1 at the i^{th} vertex and 0 at all

other vertices; w_is are the random weights that need to be estimated, and n is the number of vertices in the mesh. Based on a set of initial points representing the observed data, a typical mesh is chosen to maximize the minimum interior triangle angle to ensure smooth transitions between small and large triangles. The vertices are often chosen to be the data locations, and extra vertices are added heuristically to minimize the total number of triangles needed to fulfill the size and shape constraints of the function domain. Automated mesh construction based on a set of initial points' Euclidean coordinates in $\mathbb{R}^2$ or $\mathbb{S}^2$ is implemented in the `R-INLA` package (Martins et al. 2013). The weights $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)'$ are then the values of the field at each vertex in the mesh, and the values in the interior of the triangles are determined by linear interpolation.

The joint distribution of $\boldsymbol{w}$ is chosen so that the distribution of the functions $\beta(\boldsymbol{u})$ approximates the distribution of solutions to (2). The result is that $\boldsymbol{w}$ is Gaussian with mean $\mathbf{0}$ and sparse precision matrix $\boldsymbol{Q}_{\kappa,\tau} = \tau^2 (\kappa^4 \boldsymbol{C} + 2\kappa^2 \boldsymbol{G} + \boldsymbol{G}\boldsymbol{C}^{-1}\boldsymbol{G})$, where $\boldsymbol{G}$ is a sparse symmetric $n \times n$ matrix with non-zero entries in cells corresponding to neighboring locations and $\boldsymbol{C}$ is a diagonal matrix (Bolin and Lindgren 2013). Consider N data locations $\boldsymbol{u}_i$ $(i = 1, \ldots, N)$ and let vector $\boldsymbol{\beta}$ contain a realization of the random field at those locations. Then the SPDE prior on $\boldsymbol{\beta}$ is given by

$$\boldsymbol{\beta} = \boldsymbol{\Psi}\boldsymbol{w}, \quad \boldsymbol{w} \mid \kappa, \tau \sim N\left(\mathbf{0}, \boldsymbol{Q}_{\kappa,\tau}^{-1}\right), \tag{4}$$

where $\boldsymbol{\Psi}$ is the $N \times n$ sparse matrix of the basis functions. Note that $\boldsymbol{\Psi}$ is the identity matrix if the mesh vertices are the data locations. We take independent log-normal priors on the hyperparameters κ and τ (Lindgren and Rue 2015).

2.2 Approximate Inference by INLA

Based on model (1) and the priors specified above, we may construct a Bayesian hierarchical model for our fMRI analysis as follows

$$\begin{aligned}
&\boldsymbol{y} \mid \boldsymbol{\beta}_k, \boldsymbol{b}_j, \boldsymbol{\theta} \sim N\left(\boldsymbol{\mu}_y, \boldsymbol{V}\right), \quad \boldsymbol{\mu}_y = \sum_{k=0}^{K} \boldsymbol{X}_k \boldsymbol{\beta}_k + \sum_{j=1}^{J} \boldsymbol{Z}_j \boldsymbol{b}_j, \\
&\boldsymbol{\beta}_k = \boldsymbol{\Psi}_k \boldsymbol{w}_k, \quad \boldsymbol{w}_k \mid \boldsymbol{\theta} \sim N\left(\mathbf{0}, \boldsymbol{Q}_{\kappa_k, \tau_k}^{-1}\right), \\
&\boldsymbol{b}_j \sim N(\mathbf{0}, \delta \boldsymbol{I}), \quad \boldsymbol{\theta} \sim \pi(\boldsymbol{\theta}),
\end{aligned} \tag{5}$$

where $\boldsymbol{\theta} = (\xi, \phi_1, \ldots, \phi_p, \kappa_0, \ldots, \kappa_K, \tau_0, \ldots, \tau_K)$ and $\pi(\boldsymbol{\theta})$ denotes the joint prior density.

It is possible in theory to derive the full conditional distribution of each unknown parameter, and then use MCMC-based algorithms to obtain quite a few samples from their posterior distributions and make Bayesian inferences using those samples. However, the MCMC may have mixing problems and be slow to converge. As an alternative, Rue et al. (2009) introduced a novel Bayesian computation tool for latent Gaussian models based on integrated nested Laplace approximations (INLA), which produces very good approximations of the marginal posterior densities and computes all necessary estimates faster than sampling-based techniques. We now briefly describe the INLA method, referring the reader to Rue et al. (2009) for details.

A typical latent Gaussian hierarchical model has a set of hyperparameters $\boldsymbol{\theta}$ with prior $\pi(\boldsymbol{\theta})$, a set of latent Gaussian variables (e.g., latent fields) $\boldsymbol{f}$ with prior $\pi(\boldsymbol{f}|\boldsymbol{\theta})$ and a set of response variables $\boldsymbol{y}$ with likelihood $\pi(\boldsymbol{y}|\boldsymbol{f}, \boldsymbol{\theta})$. The joint posterior distribution is then given by $\pi(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{y}) \propto \pi(\boldsymbol{y}|\boldsymbol{f}, \boldsymbol{\theta})\pi(\boldsymbol{f}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$. The INLA method first uses Laplace approximations to approximate the marginal posterior of $\boldsymbol{\theta}$, denoted $\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})$. The approximated marginals of each θ_j can be obtained by summing out the remaining variables $\boldsymbol{\theta}_{-j}$ from $\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})$. If needed, the approximated marginal of f_i is obtained by first approximating the full conditional of f_i with another Laplace approximation. The parameters are then numerically integrated out from $\tilde{\pi}(f_i|\boldsymbol{\theta}, \boldsymbol{y})$, which gives

$$\tilde{\pi}(f_i|\boldsymbol{y}) \approx \sum_{\ell} \lambda_\ell \tilde{\pi}(f_i|\boldsymbol{y}, \boldsymbol{\theta}_\ell), \tag{6}$$

where λ_ℓ are proportional to $\tilde{\pi}(\boldsymbol{\theta}_\ell|\boldsymbol{y})$. The evaluation points $\boldsymbol{\theta}_\ell$ can be chosen in different ways,

depending on the importance of computational efficiency in a given setting (Martins et al. 2013).

For the proposed Bayesian GLM model, given $\boldsymbol{f} = (\boldsymbol{\beta}_0', \ldots, \boldsymbol{\beta}_K', \boldsymbol{b}_1', \ldots, \boldsymbol{b}_J')'$, $\pi(\boldsymbol{y}|\boldsymbol{f}, \boldsymbol{\theta})$ is the Gaussian likelihood function defined in (5); $\pi(\boldsymbol{f}|\boldsymbol{\theta})$ is the joint multivariate Gaussian distribution of the independent priors specified on $\boldsymbol{b}_j$ and $\boldsymbol{\beta}_k$; and $\pi(\boldsymbol{\theta})$ is the joint prior density of the hyper-parameters. Since the likelihood is Gaussian for this model, the only necessary approximation is the numerical integration over $\pi(\boldsymbol{\theta}|\boldsymbol{y})$.

2.3 Joint PPM for Activation Identification

After fitting model (5) with INLA, we may use the resulting estimates of activation amplitudes to identify activated brain regions. Existing thresholding techniques are based on first calculating the *marginal* probabilities $P(f(\boldsymbol{u}) > \gamma)$, where γ is an activation threshold, then defining the excedence region as $D = \{\boldsymbol{u} : P(f(\boldsymbol{u}) > \gamma) > 1-\alpha\}$, where α is some significance level. However, the value of α needs to be adjusted for multiple comparisons, which is typically done using Type I error control, false discovery rate thresholding, or posterior probability thresholding (Marchini and Presanis 2004). Here, we instead employ the *joint* posterior probabilities using the excursion set method introduced by Bolin and Lindgren (2015). This results in an estimated joint posterior probability map (PPM) that can then be thresholded at a given significance level to identify the regions that are jointly activated in response to a particular task.

Define the positive excursion set $E^+_{\gamma,\alpha}$ as the largest set of vertices such that with probability at least $1-\alpha$ the level γ is exceeded at *all* locations in that set, which we can write as

$$E^+_{\gamma,\alpha}(f) = \arg\max_D \{|D| : P(D \subseteq A^+_\gamma(f)) \geq 1-\alpha\},$$

where $A^+_\gamma(f) = \{\boldsymbol{u} \in \Omega; f(\boldsymbol{u}) > \gamma\}$. Since $E^+_{\gamma,\alpha}(f)$ is defined using the joint distribution of the random field, it should be calculated based on the posterior distribution

$$\pi(\boldsymbol{f} \mid \boldsymbol{y}) = \int \pi(\boldsymbol{f} \mid \boldsymbol{y}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta} \mid \boldsymbol{y}) d\boldsymbol{\theta}. \tag{7}$$

Computing $E^+_{\gamma,\alpha}(f)$ based on this distribution is computationally demanding, but can be done efficiently using the INLA technique. To do so, note that $E^+_{\gamma,\alpha}(f) = A^+_{1-\alpha}(F^+_\gamma)$ where $F^+_\gamma(\boldsymbol{u}) = \sup\{1-\alpha, \boldsymbol{u} \in E^+_{\gamma,\alpha}(f)\}$ is the *excursion function*. The excursion function can be estimated

as $F^+_\gamma(\boldsymbol{u}) = \sum_{\ell=1}^{L} \lambda_\ell F_\ell(\boldsymbol{u})$ where $F_\ell(\boldsymbol{u})$ is the excursion function calculated for the conditional posterior $\pi(\boldsymbol{f}|\boldsymbol{y}, \boldsymbol{\theta}_\ell)$ for a fixed parameter configuration $\boldsymbol{\theta}_\ell$ with corresponding weights λ_ℓ, as in (6).

Here, $\pi(\boldsymbol{f}|\boldsymbol{y}, \boldsymbol{\theta}_\ell)$ is Gaussian and the computation of $F_\ell(\boldsymbol{u})$, $\ell = 1, \ldots, L$ therefore only requires the ability to compute excursion probabilities of multivariate Gaussian distributions. This can be done efficiently using the sequential method described in Bolin and Lindgren (2015). We refer to Bolin and Lindgren (2015) for further details and note that the method is implemented in the R package `excursions` (Bolin and Lindgren 2018), which has an interface to `R-INLA`.

3 MULTI-SUBJECT BAYESIAN GLM

So far, we have been concerned with the analysis of a single subject's data. In practice, however, researchers often want to estimate population-level effects or pool information across one or more groups of subjects. While it is common in the classical GLM context to use a hierarchical group model (e.g., Worsley et al. 2002, Woolrich et al. 2004, Lindquist et al. 2012), previously proposed spatial Bayesian GLMs have been limited to single-subject analysis due to computational challenges. Currently there exists no spatial Bayesian modeling approach for task fMRI analysis that is applicable to multi-subject studies, something that is highly desirable for the neuroimaging community (Sidén et al. 2017). Consider a second-level model for L group-level activations or contrasts of interest, given by

$$\boldsymbol{\beta} = \boldsymbol{X}_G \boldsymbol{\beta}_G + \boldsymbol{\varepsilon}_G, \quad \boldsymbol{\varepsilon}_G \sim N\left(\boldsymbol{0}, \sigma^2_G \boldsymbol{I}\right), \tag{8}$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}'_{11}, \ldots, \boldsymbol{\beta}'_{1K}, \ldots, \boldsymbol{\beta}'_{M1}, \ldots, \boldsymbol{\beta}'_{MK})'$; $\boldsymbol{X}_G$ is a $NKM \times NL$ second-level design matrix; $\boldsymbol{\beta}_G$ is the $NL \times 1$ vector of group-level features; and σ^2_G is the between-subject variance. For example, in the case that we are interested in the average activation for each task for a single group of subjects with no covariate effects, we have $\boldsymbol{X}_G = \boldsymbol{1}_M \otimes \boldsymbol{I}_{NK}$.

One naive solution would be to first fit M subject-level models to obtain the posterior mean of $\boldsymbol{\beta}$, denoted $\hat{\boldsymbol{\beta}}$, then use $\hat{\boldsymbol{\beta}}$ as the response variable in equation (8). Taking SPDE priors on the components of $\boldsymbol{\beta}_G$, computation can be performed using INLA as described in Section 2.2, and inference on $\boldsymbol{\beta}_G$ can be performed as described in Section 2.3. However, assuming SPDE priors

on both the subject-level and population-level activation fields may lead to over-smoothing of the group effects. Such a two-level approach is therefore not ideal.

Given the computational advantages of the proposed methods (see Section 5), it may in some cases be feasible to include data from multiple subjects in a single model to estimate group effects of interest. We describe such a "fully Bayesian" approach next. However, in many cases, such as large group studies, fitting such a model will be computationally prohibitive. Therefore, we also propose a novel approach to combining the results of the subject-level models in a principled manner to obtain posterior estimates of the group effects.

3.1 Fully Bayesian Modeling Approaches

For simplicity, consider without loss of generality the case with no nuisance covariates and no baseline field, as these can be regressed out of the fMRI response and task design for each subject a-priori. A natural approach would be to estimate group effects directly in a multi-subject model, with subject-level deviations modeled as random effects. However, estimating such a model is very computationally demanding, given the increased number of latent fields and hyperparameters in the model. A more computationally efficient approach would be to model group effects alone, ignoring subject-level deviations. For example, if we wish to estimate the average activation across subjects for each task, we can consider the group-effects (GE) model

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{X}_k \boldsymbol{\beta}_k^{pop} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \boldsymbol{I}_m \otimes \boldsymbol{V}), \tag{9}$$

where $\boldsymbol{y}_m$ is the fMRI time series for subject m, $\boldsymbol{X}_{mk}$ is their task design matrix for task k, $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_M')'$ and $\boldsymbol{X}_k = (\boldsymbol{X}_{1k}', \ldots, \boldsymbol{X}_{Mk}')'$. We assume an SPDE prior on each $\boldsymbol{\beta}_k^{pop}$ with SPDE parameters κ_k^{pop} and τ_k^{pop}. Although this model ignores the effect of subject-level deviations on the covariance structure, through simulations (results not shown) we found it to have very similar performance to a random effects model and, as reported in Section 4.2 below, to be quite accurate in estimating group effects and identifying areas of activation.

Alternatively, since subject-level activation fields are often themselves of interest, they can be modeled as fixed effects, with group effects being modeled as linear combinations of the

subject-level effects. Consider the following fixed effects (FE) model

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{X}_k \boldsymbol{\beta}_k + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \boldsymbol{I}_m \otimes \boldsymbol{V}), \tag{10}$$

where $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_M')'$, $\boldsymbol{X}_k = \text{diag}\{\boldsymbol{X}_{1k}, \ldots, \boldsymbol{X}_{Mk}\}$, and $\boldsymbol{\beta}_k = (\boldsymbol{\beta}_{1k}', \ldots, \boldsymbol{\beta}_{Mk}')'$. We assume an SPDE prior on each $\boldsymbol{\beta}_{mk}$ with SPDE parameters κ_k and τ_k common to all subjects. The residual precision structure $\boldsymbol{V}$, specified in the beginning of this section, is also common across subjects.

Letting $\boldsymbol{\beta} = (\boldsymbol{\beta}_{11}', \ldots, \boldsymbol{\beta}_{1K}', \ldots, \boldsymbol{\beta}_{M1}', \ldots, \boldsymbol{\beta}_{MK}')'$ be the vector of all subject-level activation fields, we define the group-level $\boldsymbol{\beta}_G$ to be a linear combination of the subject-level activations, i.e., $\boldsymbol{\beta}_G = \boldsymbol{A}\boldsymbol{\beta}$, where $\boldsymbol{A}$ is a constant matrix determined by the nature of $\boldsymbol{\beta}_G$. For example, again considering the average activation of the M subjects for each task, we may define $\boldsymbol{A} = \frac{1}{M}\boldsymbol{1}_M' \otimes \boldsymbol{I}_{NK}$, so that $\boldsymbol{\beta}_G = \sum_{m=1}^{M} \boldsymbol{\beta}_m / M$. $\boldsymbol{A}$ can also be defined to estimate contrasts between tasks, group differences, or covariate effects.

Given the estimated posterior mean of $\boldsymbol{\beta}$, the posterior mean of $\boldsymbol{\beta}_G$ can be easily estimated. For identification of areas of activation, note that the posterior distribution of $\boldsymbol{\beta}$ is Gaussian for a given hyperparameter configuration. We can therefore draw samples from the posterior of $\boldsymbol{\theta}$, then for each such sample, draw Monte Carlo samples from the conditional posterior of $\boldsymbol{\beta}$, and finally compute $\boldsymbol{\beta}_G$ for each sampled $\boldsymbol{\beta}$. We can then utilize the `excursions.mc` function in the `excursions` R package to estimate the excursion function for each group-level activation field, which can then be thresholded to identify areas of activation.

3.2 Joint Modeling Approach

We now present an alternative approach to group-level inference based on fitting each subject-level model separately then combining the results in a principled manner. We call this the "joint" modeling approach, since it is based on the joint posterior distribution of subject-level activation fields. Combining across tasks $k = 1, \ldots, K$, let $\boldsymbol{\beta}_m = (\boldsymbol{\beta}_{m1}', \ldots, \boldsymbol{\beta}_{mK}')'$ for each subject m. Using the notation of equation (4), $\boldsymbol{\beta}_m = \boldsymbol{\Psi}_m \boldsymbol{w}_m$, where $\boldsymbol{\Psi}_m = \text{diag}(\boldsymbol{\Psi}_{m1}, \ldots, \boldsymbol{\Psi}_{mK})$ and $\boldsymbol{w}_m = (\boldsymbol{w}_{m1}', \ldots, \boldsymbol{w}_{mK}')'$. Taking an SPDE prior on each $\boldsymbol{\beta}_{mk}$ implies that $\boldsymbol{w}_m \mid \boldsymbol{\theta} \overset{i.i.d.}{\sim} N\left(\boldsymbol{0}, \boldsymbol{Q}_\theta^{-1}\right)$ with $\boldsymbol{Q}_\theta = \text{diag}(\boldsymbol{Q}_{\kappa_1,\tau_1}, \ldots, \boldsymbol{Q}_{\kappa_K,\tau_K})$, where $\boldsymbol{Q}_{\kappa_k,\tau_k}$ denotes the precision matrix of the SPDE prior

taken on $\boldsymbol{\beta}_{mk}$. Then, the full conditional distribution of $\boldsymbol{w}_m$ is given by

$$\boldsymbol{w}_m \mid \boldsymbol{y}_m, \boldsymbol{\theta} \sim N\left(\boldsymbol{\mu}_m, \boldsymbol{Q}_m^{-1}\right), \tag{11}$$

where $\boldsymbol{\mu}_m = \boldsymbol{Q}_m^{-1}\boldsymbol{\Psi}'_m\boldsymbol{X}'_m\boldsymbol{V}^{-1}\boldsymbol{y}_m$ and $\boldsymbol{Q}_m = \boldsymbol{Q}_\theta + \boldsymbol{\Psi}'_m\boldsymbol{X}'_m\boldsymbol{V}^{-1}\boldsymbol{X}_m\boldsymbol{\Psi}_m$. Since subjects are assumed to be independent, the full conditional distribution of $\boldsymbol{w} = (\boldsymbol{w}'_1, \ldots, \boldsymbol{w}'_M)'$ is Gaussian with mean $\boldsymbol{\mu} = (\boldsymbol{\mu}'_1, \ldots, \boldsymbol{\mu}'_M)'$ and precision matrix $\boldsymbol{Q} = \text{diag}(\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_M)$. Letting $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \ldots, \boldsymbol{\beta}'_M)'$, we have $\boldsymbol{\beta} = \boldsymbol{\Psi}\boldsymbol{w}$, where $\boldsymbol{\Psi} = \text{diag}(\boldsymbol{\Psi}_1, \ldots, \boldsymbol{\Psi}_M)$.

We again define the group-level $\boldsymbol{\beta}_G$ to be a linear combination of the subject-level $\boldsymbol{\beta}$, i.e., $\boldsymbol{\beta}_G = \boldsymbol{A}\boldsymbol{\beta}$. Since $\boldsymbol{\beta}_G = \boldsymbol{A}\boldsymbol{\Psi}\boldsymbol{w}$ and the full conditional of $\boldsymbol{w}$ is Gaussian, it is easy to see that the full conditional of $\boldsymbol{\beta}_G$ is also Gaussian, namely $\boldsymbol{\beta}_G \mid \boldsymbol{y}, \boldsymbol{\theta} \sim N\left(\boldsymbol{A}\boldsymbol{\Psi}\boldsymbol{\mu}, \boldsymbol{A}\boldsymbol{\Psi}\boldsymbol{Q}^{-1}\boldsymbol{\Psi}'\boldsymbol{A}'\right)$. To obtain the posterior density of $\boldsymbol{\beta}_G$, we need to integrate out $\boldsymbol{\theta}$, i.e.,

$$\pi(\boldsymbol{\beta}_G \mid \boldsymbol{y}) = \int \pi(\boldsymbol{\beta}_G \mid \boldsymbol{y}, \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{y})d\boldsymbol{\theta}.$$

Unfortunately, it is hard (if not impossible) to explicitly solve the integral above. We therefore propose to numerically evaluate it using importance sampling. The marginal posterior density is $\pi(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{y} \mid \boldsymbol{\theta})$, and since the subjects are assumed to be independent given the parameters, it is straightforward to show that

$$\pi(\boldsymbol{y} \mid \boldsymbol{\theta}) \propto \pi(\boldsymbol{\theta})^{-M} \prod_{m=1}^{M} \pi(\boldsymbol{\theta} \mid \boldsymbol{y}_m), \tag{12}$$

where $\pi(\boldsymbol{\theta})$ is the prior density of $\boldsymbol{\theta}$ as specified in Section 2, and $\pi(\boldsymbol{\theta} \mid \boldsymbol{y}_m)$ is the posterior density of $\boldsymbol{\theta}$ for the m^{th} subject. The proof for (12) is given in Appendix A. We approximate $\pi(\boldsymbol{\theta} \mid \boldsymbol{y}_m)$ using the Gaussian approximation with mean $\tilde{\boldsymbol{\mu}}_m$ and precision $\tilde{\boldsymbol{Q}}_m$ given by the INLA method when fitting the subject-level models. Then, the product term in (12) is proportional to another Gaussian density, denoted by $q(\boldsymbol{\theta}|\boldsymbol{y})$, with precision matrix $\sum_m \tilde{\boldsymbol{Q}}_m$ and mean $(\sum_m \tilde{\boldsymbol{Q}}_m)^{-1} \sum_m \tilde{\boldsymbol{Q}}_m\tilde{\boldsymbol{\mu}}_m$. As a result, the posterior distribution of $\boldsymbol{\beta}_G$ can be approximated by

$$\tilde{\pi}(\boldsymbol{\beta}_G \mid \boldsymbol{y}) = \int \pi(\boldsymbol{\beta}_G \mid \boldsymbol{y}, \boldsymbol{\theta})\pi(\boldsymbol{\theta})^{1-M} q(\boldsymbol{\theta} \mid \boldsymbol{y})d\boldsymbol{\theta}.$$

Let $\ell = 1, \ldots, L$ index samples $\boldsymbol{\theta}^{(\ell)}$ drawn from $q(\boldsymbol{\theta}|\boldsymbol{y})$. Then, $\tilde{\pi}(\boldsymbol{\beta}_G \mid \boldsymbol{y})$ can be approximated numerically as a mixture of Gaussians with weights $\lambda_\ell \propto \pi(\boldsymbol{\theta}^{(\ell)})^{1-M}$, $\sum_\ell \lambda_\ell = 1$, and the posterior quantities of $\boldsymbol{\beta}_G$ can be easily computed. It's noteworthy that the highly sparse $\boldsymbol{Q}$ makes the computation efficient in spite of its large dimension. We can likewise approximate the joint PPMs as described in Section 2.3. Specifically, the excursion function F_γ for the joint posterior $\tilde{\pi}(\boldsymbol{\beta}_G \mid \boldsymbol{y})$ can be written as $F_\gamma = \sum_{\ell=1}^{L} \lambda_\ell F_\gamma^{(\ell)}$, where $F_\gamma^{(\ell)}$ is the excursion function based on the Gaussian distribution $\pi(\boldsymbol{\beta}_G \mid \boldsymbol{y}, \boldsymbol{\theta}^{(\ell)})$, which can be computed as described in Section 2.3.

In the next section, we use simulated data to assess the performance of the fully Bayesian GE and FE models and the joint modeling approach. We also compare with the naive two-level approach discussed at the beginning of this section as a baseline.

4 SIMULATION STUDY

We perform two simulation studies to assess the performance of the proposed methods for single-subject and multi-subject analysis. In the first, we compare the proposed single-subject Bayesian GLM with the classical GLM. In the second, we perform multi-subject analysis using the fully Bayesian GE and FE models, the proposed joint approach, and the naive two-level approach.

4.1 Single-Subject Simulation Study

We constructed a 46×55 phantom image using a binary brain mask provided by the SPM8 software package (4667, 27 Feb 2012; Wellcome Trust, UCL), which contains 1256 voxels within the brain. The field-of-view of the image was assumed to be 192 mm. A dynamic image time series of length $T = 200$ was simulated as follows. Two task activation profiles $x_1(t)$ and $x_2(t)$, depicted in the first column of Figure 3, were created by convolving a canonical hemodynamic response function (Friston et al. 1998) with two task-specific binary stimulus functions. Activation amplitude maps $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ were formed by placing Dirac functions at $(12, 28)$ (region 1), $(36, 28)$ (region 2) and $(23, 16)$ (region 3) and smoothing each with a Gaussian kernel with full width at half maximum (FWHM) of 15, 20 and 25 mm, respectively.

The data at locations $v = 1, \ldots, 1256$ and time points $t = 1, \ldots, 200$ was simulated as $y_v(t) = \beta_{0v} + a_{1v}x_1(t)\beta_{1v} + a_{2v}x_2(t)\beta_{2v} + \varepsilon_v(t)$, where β_{0v} is a baseline field based on the gray matter probability mask; and $\varepsilon_v(t)$ is spatially independent AR(1) errors with standard deviation

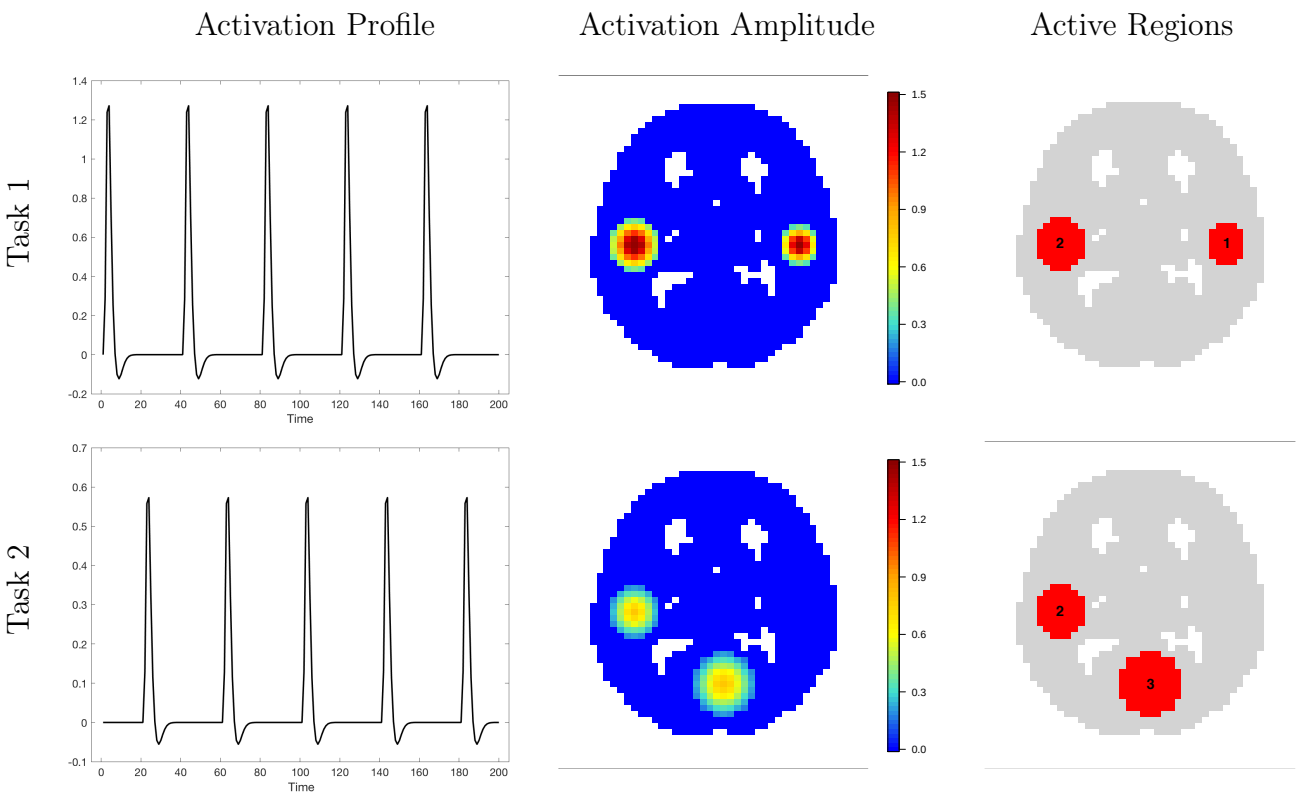

Figure 3: Activation profiles (left), activation amplitude maps in units of percent signal change (middle), and active regions based on activation threshold $\gamma = 0$ (right) for both tasks in the simulation study. Task 1 activates regions 1 and 2; task 2 activates regions 2 and 3. Activation amplitudes are in units of percent local signal change.

of 2 and autocorrelation coefficient of 0.3. The weights (a_{1v}, a_{2v}) were $(4, 0)$ within region 1, $(4, 2)$ within the region 2, and $(0, 2)$ within region 3. The resulting activation amplitudes and activated regions are shown in Figure 3.

Prior to model fitting, the simulated fMRI data was scaled to units of percent local signal change as $\tilde{y}_v(t) = (y_v(t) - \bar{y}_v)/\bar{y}_v$, where $\bar{y}_v$ is the mean fMRI signal at voxel v, so that coefficient estimates represent percent signal change in response to each task. We fit the model

$$\tilde{y}_v(t) = \tilde{\beta}_{0v} + \tilde{\beta}_{1v} x_1(t) + \tilde{\beta}_{2v} x_2(t) + \tilde{\varepsilon}_v(t) \tag{13}$$

with the priors given in equation (5) using the proposed Bayesian GLM approach. Figure 4 displays the triangular mesh, where we see regular triangulation inside the brain and two boundary layers of larger triangles surrounding the brain and within the ventricles. Areas of activation were identified using the joint PPM approach with activation threshold $\gamma = 0$. Bayesian GLM model

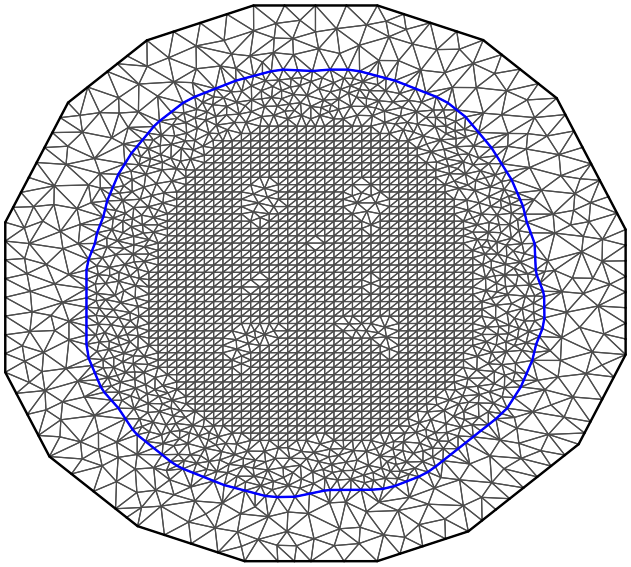

Figure 4: The triangular mesh used in the simulation study. We see regular triangulation inside the brain and two boundary layers of larger triangles surrounding the brain and within the ventricles. This encourages smoothing along the boundaries of the brain without blurring the boundaries.

fitting was done by implementing the `R-INLA` package with the `PARDISO` sparse matrix library and 4 parallel threads on a server with 64GB of memory. Computation took approximately 25 minutes, plus 10 minutes per task to estimate the joint PPM using the `excursions` package.

For comparison, we applied the classical GLM approach to the simulated data both with and without spatial smoothing. For the spatially smoothed data, we first convolved each fMRI image with a 2-dimensional 6 mm FWHM Gaussian kernel. We then prewhitened the data to remove autocorrelation of the residuals by fitting a linear regression model assuming independent errors to each voxel separately, estimating the AR coefficient at each voxel by solving the Yule-Walker equations of the model residuals, then averaging across voxels. Regions of activation were identified by performing a t-test at every voxel, correcting for multiple comparisons with FDR control using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995) and FWER control by performing a permutation test based on 1000 random reorderings of the prewhitened time series.

Estimates of the two amplitude fields based on the Bayesian and classical GLMs are shown in Figure 5. All three sets of estimates capture the general spatial patterns of the true fields, but the classical estimates are much more noisy, particularly when the data is not smoothed.

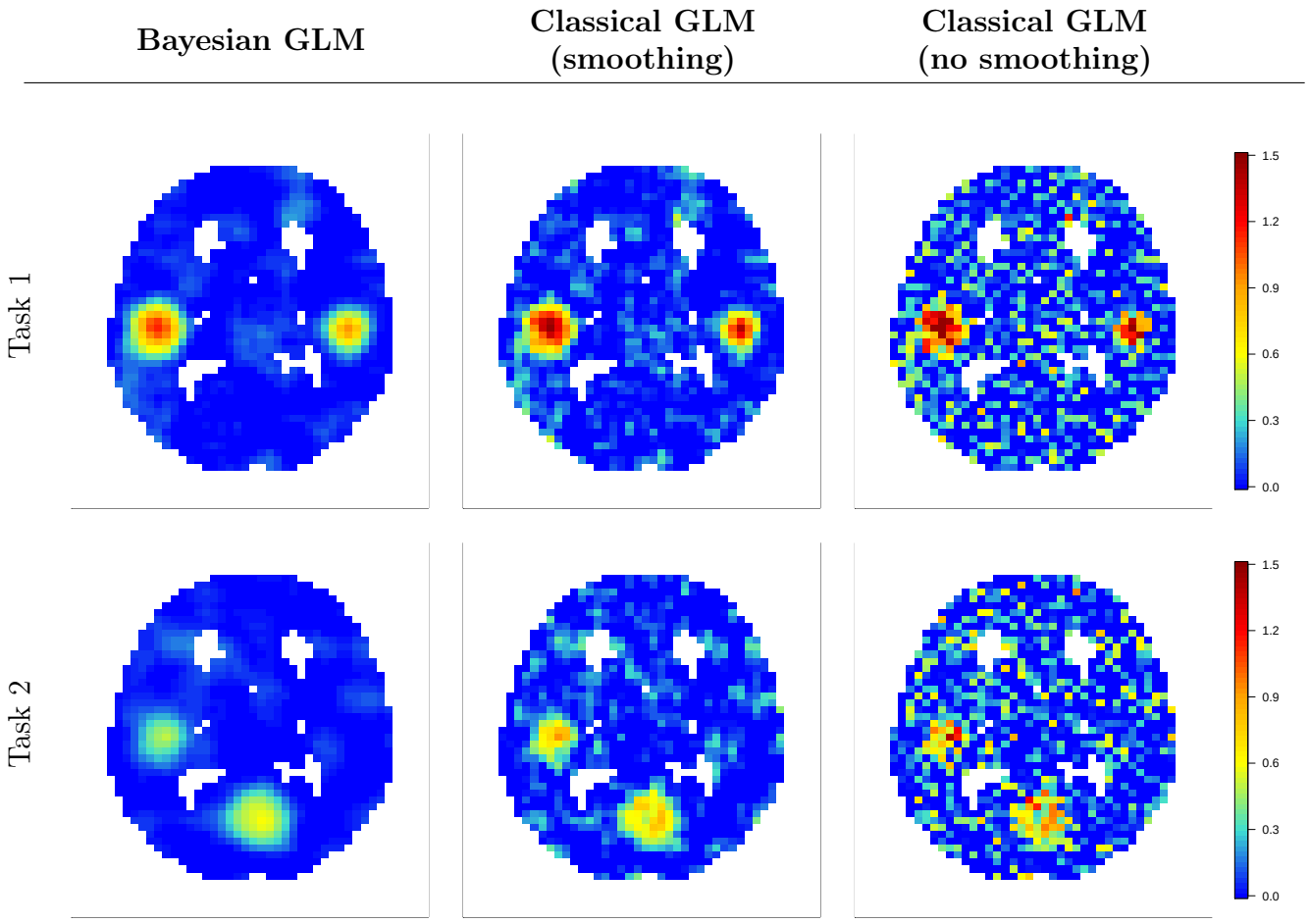

Figure 5: Bayesian and classical estimates of activation amplitudes in simulation study, in units of percent local signal change.

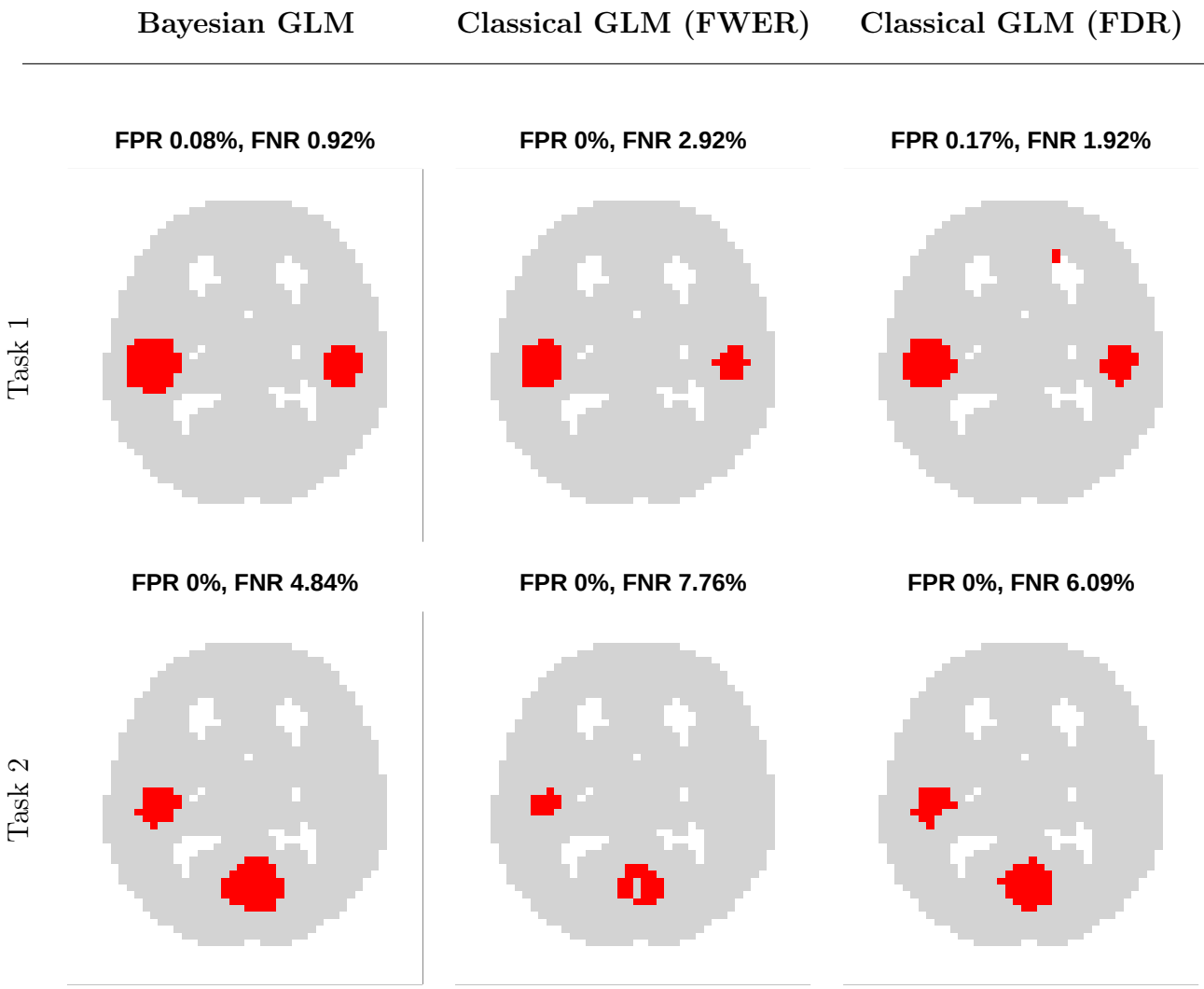

Figure 6: Estimated regions of activation in the simulation study. The false positive rate (FPR) and false negative rate (FNR) are reported for each method and activation. In the Bayesian GLM, regions of activation are estimated using the joint PPM approach with $\alpha = 0.01$ and $\gamma = 0\%$ signal change. In the classical GLM, regions of activation are estimated by performing a hypothesis test on the task coefficient at each location, correcting for multiple comparisons through FDR control ($q = 0.01$) and FWER control ($\alpha = 0.01$). Classical GLM results are based on smoothed data; results based on unsmoothed data are given in Appendix B.

The Bayesian GLM tends to somewhat underestimate the peak of activation as a result of the assumption of stationarity, which assumes that the smoothness around areas of activation is the same as the smoothness of background regions. The classical GLM with smoothing also results in slightly attenuated activation amplitudes. Figure 6 displays estimated regions of activation based on each method. The regions of activation based on the Bayesian joint PPM method appear most similar to the true regions of activation. FWER control and to a lesser degree FDR control yields a higher rate of false negatives, suggesting that that the Bayesian approach has higher power to detect true effects.

Figure 7 displays receiver operating characteristic (ROC) curves for the Bayesian and classical GLMs. The ROC curves illustrate that the Bayesian approach achieves both high sensitivity and specificity, with 0.9993 and 0.9869 area under the curve (AUC) for tasks 1 and 2, respectively. In contrast, the classical GLM is not able to simultaneously achieve the same level of sensitivity and specificity, with AUC of 0.9882 and 0.9160 with smoothing or 0.9026 and 0.8033 without smoothing. Notably, the gain in performance of the Bayesian GLM over the classical GLM is greatest when specificity is high. As high specificity is often a principal goal in fMRI task activation studies, this suggests that the proposed Bayesian approach is able to achieve higher power than the classical GLM when the false positive rate is controlled at a fixed level.

The bottom row of Figure 7 compares the excursion function values obtained from the joint and marginal posterior probabilities. The points show that the excursion function based on the joint posterior probabilities is a monotonic transformation of the excursion function based on the marginal posterior probabilities. This is due to the choice of parametric family implemented in the `excursions` R package, which has the monotonic property (Bolin and Lindgren 2015). However, the distribution of values obtained from the two approaches is very different. The excursion function values based on the joint PPM (histogram shown in blue) are clustered around 0 and 1, while the values based on the marginal PPM (histogram shown in red) are continuously distributed between 0 and 1. Therefore, besides providing more accurate control of false positives, the joint PPM will also tend to be less sensitive to the choice of significance threshold.

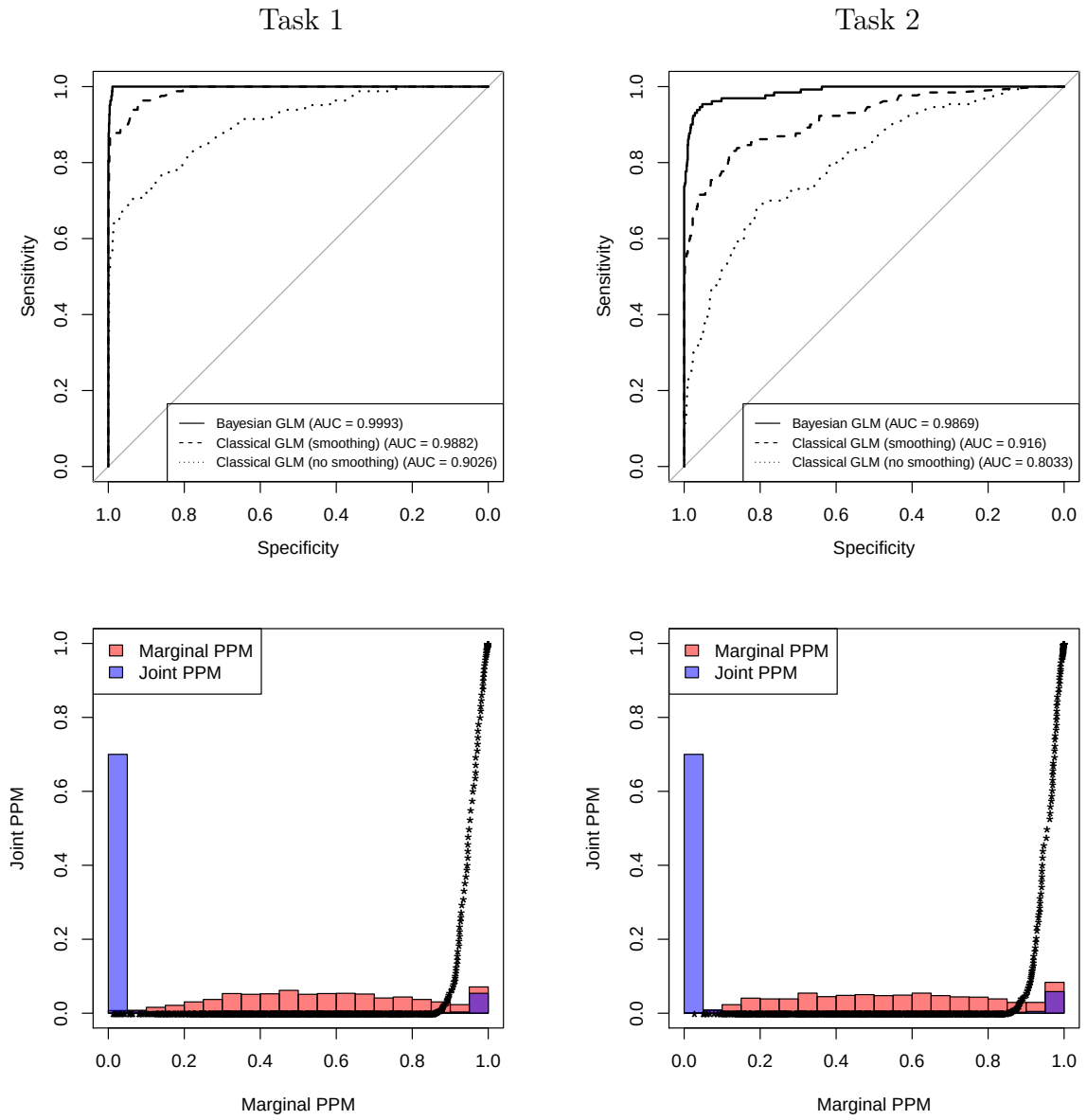

Figure 7: **Top row:** ROC curves for both tasks of the simulation study. In the Bayesian GLM, regions of activation are estimated by thresholding the excursion function at a certain level; in the classical GLM, the threshold is determined by performing a hypothesis test at each location. Since we performed voxel-level multiple comparisons correction in the classical GLM, FWER and FDR control have identical ROC curves. **Bottom row:** Excursion function values obtained from marginal and joint PPM approaches. The curves show that the excursion function based on the joint PPM is a monotonic transformation of that based on the marginal PPM, resulting in identical ROC curves. However, the distribution of excursion function values obtained from the two approaches is very different. The joint PPM results in values close to 0 and 1, and hence is less sensitive to the choice of significance threshold, while the marginal PPM results in a more continuous distribution of values between 0 and 1.

	Initial Value	Deviation
Location (voxels)	$(12, 28)$, $(36, 28)$, $(23, 16)$	independent $[N(0, \sigma = 0.5)]$
Amplitude	1, 1, 1	independent $\lvert N(0, \sigma = 1)\rvert$
Smoothness (FWHM)	15, 20, 25	independent $N(0, \sigma = 1)$

Table 1: In the multi-subject simulation, subject-level activation amplitude maps were obtained by varying the location, smoothness and intensity of each active region.

4.2 Multi-Subject Simulation Study

For the multi-subject simulation, we simulated data from 10 subjects as described in Section 4.1, except that the activation amplitude maps $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ were allowed to differ across subjects by perturbing the location, smoothness and intensity of the active regions (see Table 1). To minimize computational burden of the fully Bayesian approach, independent errors with standard deviation of 2 were generated, and the data and task timecourses were centered in order to exclude the baseline signal from the model. The true group-level activations, defined to be the average over subjects, and areas of activation based on $\gamma = 0$ are shown in Figures 8 and 10.

The fully Bayesian GE and FE models (based on 10,000 Monte Carlo samples), the proposed joint approach (based on $L = 50$) and the naive two-level approach were applied to the simulated data to estimate the group-level activation amplitudes. The joint PPM method was applied to each set of results to obtain areas of activation using an activation threshold of $\gamma = 0$ and significance level $\alpha = 0.01$. Computation was performed as described in Section 4.1.

The computation time for the fully Bayesian multi-subject GE model was approximately 30 minutes for model fitting followed by 10 minutes for excursion function estimation. For the FE model, computation time was 1.7 hours for model fitting plus 25 minutes for posterior sampling and excursion function estimation. For the joint and two-level approaches, estimating each subject-level model required 2-2.5 minutes of computation time, followed by 15 minutes for computing the group-level posterior quantities in the joint approach, or 6 minutes for group-level model fitting and excursion function estimation in the two-level approach.

The estimated activation amplitudes and areas of activation based on each method are shown in Figures 9 and 11. The fully Bayes GE model is most accurate and smoothes the areas of activation less than the FE model, joint approach or two-level approach. It also achieves the lowest FPR, at the expense of slightly more false negatives. The fully Bayes FE model and

the joint approach appear to have comparable performance, while the two-level approach is slightly less accurate overall, performs more oversmoothing than the other methods, and results in substantially higher rates of false positives.

These results suggest that the fully Bayes GE model is the most accurate in estimating group-level activations and should be used for that purpose when computationally feasible. However, the proposed joint approach also has good performance and is comparable to the fully Bayes FE model. The joint approach is computationally advantageous, as subject-level models can be estimated in parallel, and therefore presents a valuable alternative when sample sizes are too large to fit the GE model. The two-level approach should be avoided due to its tendency to oversmooth, resulting in attenuated activation amplitudes and higher rates of false positives.

5 EXPERIMENTAL DATA RESULTS

5.1 Data Collection and Processing

We applied the proposed methods to fMRI data from 20 randomly sampled subjects from the Human Connectome Project (HCP), a database of demographic, behavioral and neuroimaging data collected on over one thousand healthy adult subjects (`http://humanconnectome.org`). We employed the fully preprocessed cs-fMRI data available in the HCP 500-subject data release. We refer the reader to Van Essen et al. (2013) and Glasser et al. (2013) for detailed descriptions of scanning and processing protocols. The spherical surface representation was used for model fitting, but for display purposes we used lesser degrees of inflation to allow for visualization of basic brain shape and structure.

We analyzed the HCP motor task and gambling task studies. In the motor study, during each subject's 3.5-minute fMRI run consisting of 284 volumes, subjects performed a series of five motor tasks, each preceded by a three-second visual cue (HCP 2016). In the gambling study, during each subject's 3-minute fMRI run consisting of 253 volumes, subjects were asked to guess whether the number on a mystery card, represented by a "?", was more or less than 5. Depending on their response, the subject could win or lose money. Three types of trials—reward, neutral, and loss trials—occurred throughout each run, for a total of 32 trials.

For both studies, the fMRI data for each subject was centered and scaled by the average BOLD signal at each vertex. The coefficient estimates at each vertex therefore represent the

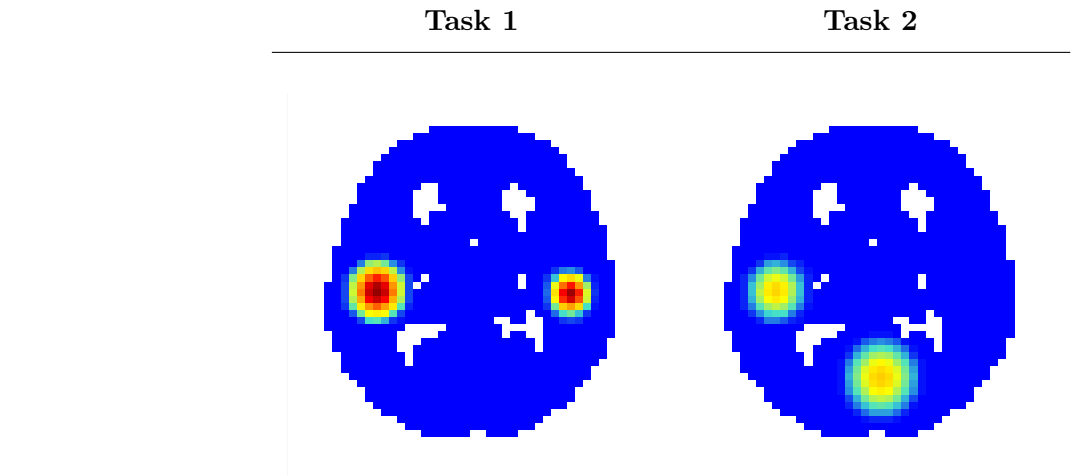

Figure 8: True simulation group-level activation amplitudes, in units of percent local signal change.

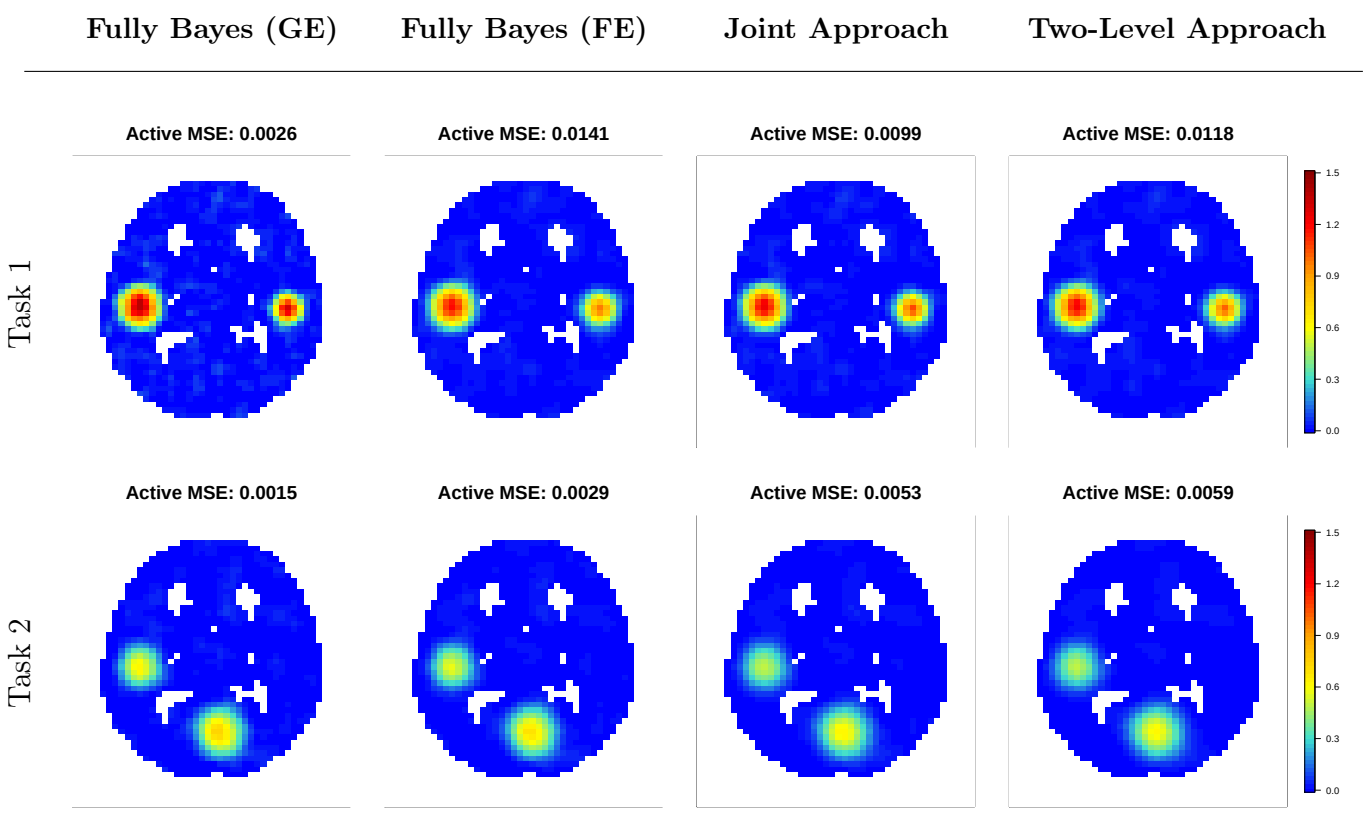

Figure 9: Estimated simulation group-level activation amplitudes, in units of percent local signal change. The mean squared area (MSE) within the true area of activation ("active MSE") is reported above each estimated map.

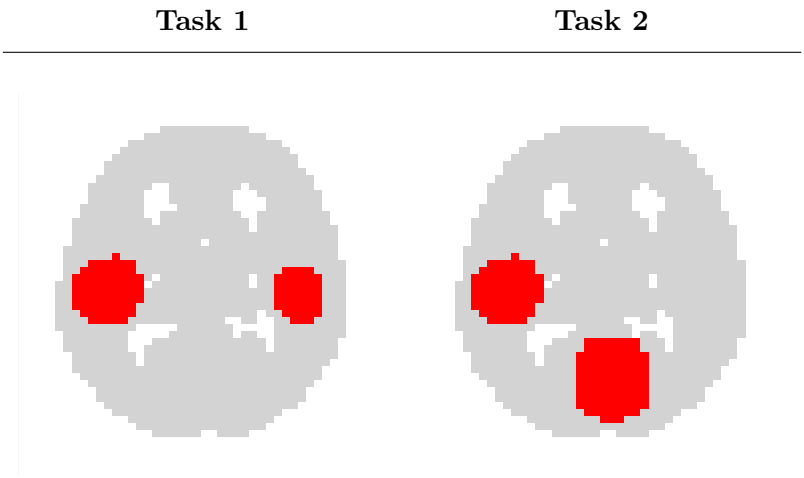

Figure 10: True simulation group-level areas of activation, based on activation threshold $\gamma = 0$.

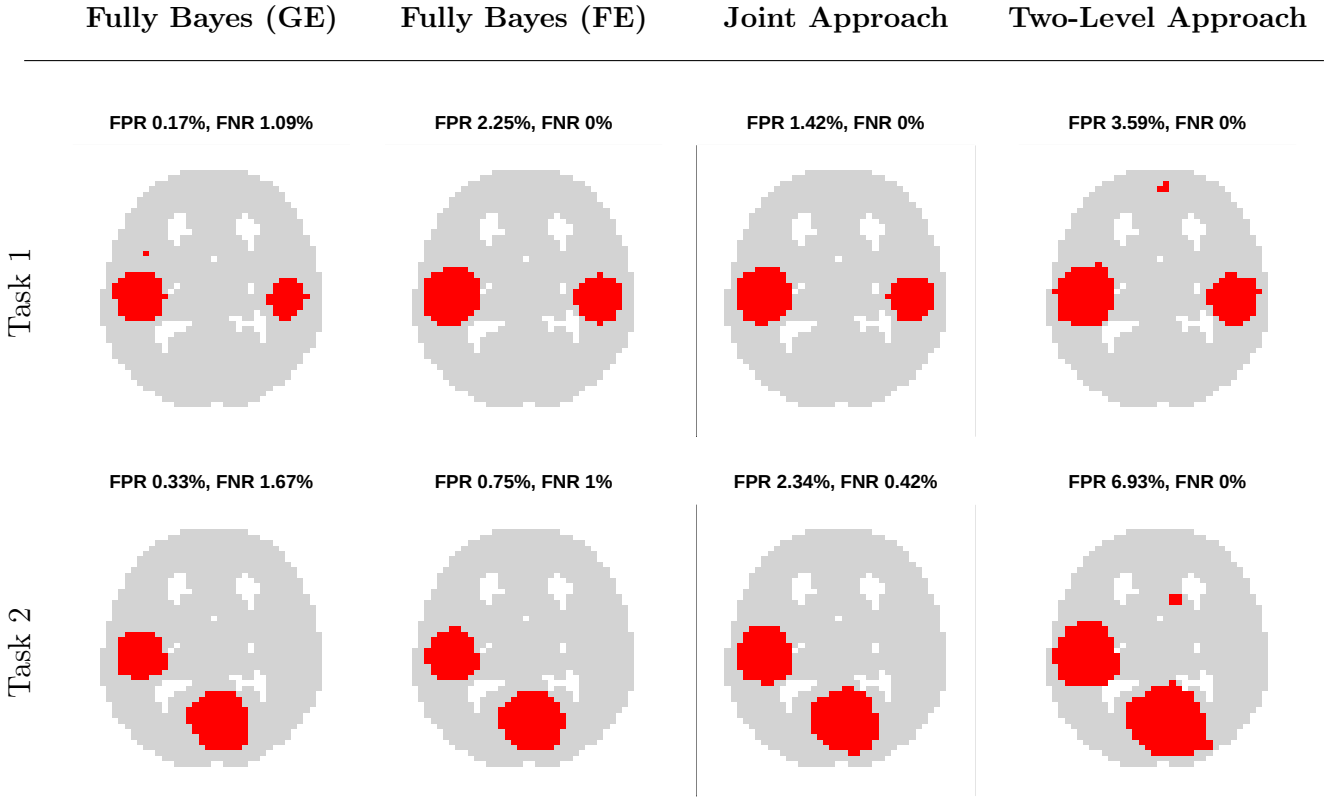

Figure 11: Estimated simulation group-level areas of activation, based on activation threshold $\gamma = 0$ and significance level $\alpha = 0.01$.

percent change in local BOLD activity due to each task. The stimulus response functions were also centered to allow removal of the intercept from the model (which is typically not of interest) and thus reduce the computational burden of the Bayesian GLM by eliminating the two hyper-parameters associated with the baseline field. To account for noise due to subject motion and scanner drift, six rigid body realignment parameters estimated in the motion realignment phase of preprocessing and their first-order temporal derivatives, along with linear and quadratic time terms, were used as nuisance covariates. To reduce the computational cost of estimating the Bayesian model, we regressed out these 14 nuisance covariates prior to model fitting.

Finally, the fMRI time courses were pre-whitened by assuming an AR(p) process on the residuals from a classical GLM with uncorrelated errors. A model order of $p = 6$ was chosen based on inspection of the partial autocorrelation functions. The AR coefficients were estimated by solving the Yule-Walker equations and were allowed to vary spatially, as previous studies have shown that the degree of temporal autocorrelation is not constant across the brain (Worsley et al. 2002, Eklund et al. 2012). We indeed observed substantial differences in the AR coefficients across the brain (Appendix Figure C3). To improve estimation efficiency, the AR coefficient estimates at each location were averaged across all 20 subjects. Investigation of residuals after prewhitening showed the remaining residual autocorrelation to be very small (Appendix Figure C3).

5.2 Model Estimation

The left and right hemisphere each contained 32,492 surface vertices. To reduce the computational burden, we used the Connectome Workbench to resample each hemisphere to approximately 6,000 vertices using barycentric interpolation to minimize smoothing (Marcus et al. 2011, Glasser et al. 2013). The average spacing between the resulting vertices was 5.0 mm (range: 4.5-5.6 mm), compared with 2.1 mm (range: 1.9-2.4 mm) prior to resampling. While this represents a decrease in spatial resolution, the surface area of each resampled vertex remains much smaller than the size of the regions of activation due to each task. The loss of information is also relatively minor compared with the common practice of spatial smoothing (Appendix Figure C4). A spherical mesh was then created based on the radial coordinates of the locations (Figure 12). The area of the medial wall, which contains missing values, is filled in with larger triangles. This has the effect of smoothing vertices along the boundary without blurring the boundary.

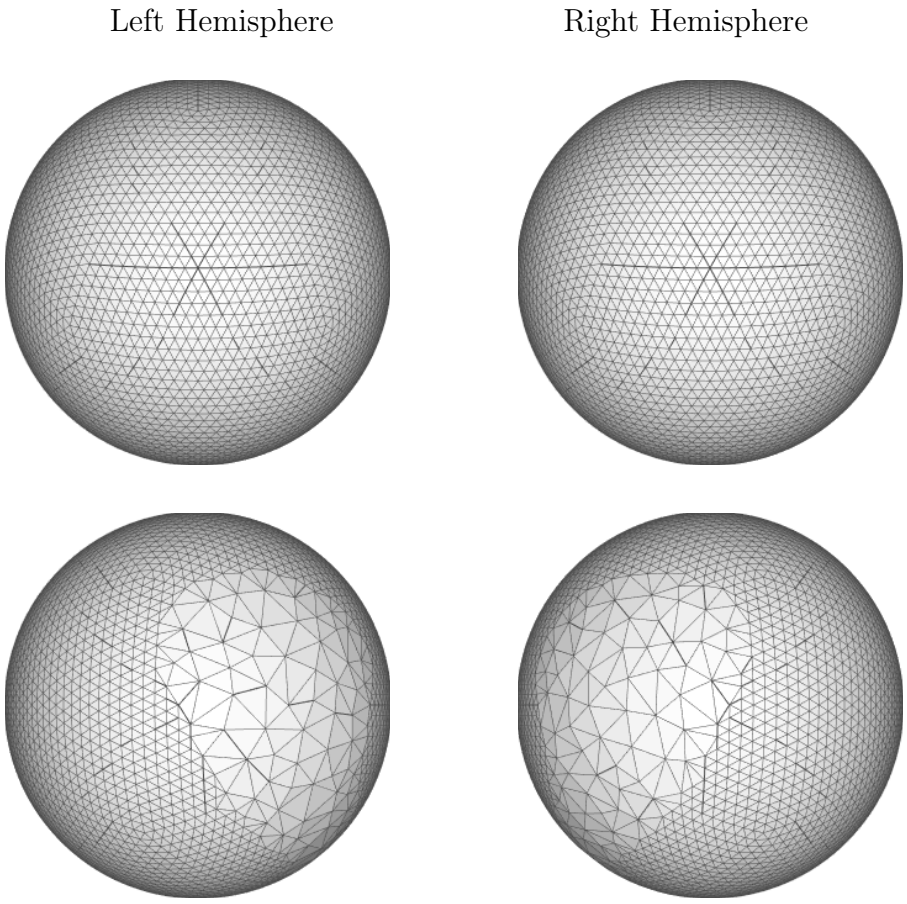

Figure 12: The spherical mesh for each hemisphere. For each hemisphere, the lateral (exterior) cortical surface is displayed on top and the medial (interior) cortical surface is displayed on bottom. The missing values in the medial wall (bottom row) are filled with larger triangles, which encourages smoothing along the boundary of the hole without blurring the boundary.

For the m^{th} ($m = 1, \ldots, 20$) subject, we estimated the following model for each hemisphere:

$$\boldsymbol{y}_m = \sum_{k=1}^{K} \boldsymbol{X}_{mk}\boldsymbol{\beta}_{mk} + \boldsymbol{\varepsilon}_m, \quad \boldsymbol{\varepsilon}_m \sim N(\boldsymbol{0}, \sigma_m^2 \boldsymbol{I}), \tag{14}$$

where $\boldsymbol{y}_m$ is a vector containing the NT response values, T is the number of time points, N is the number of locations in the mesh, and the $\boldsymbol{X}_{mk}$ are the stimulus response functions corresponding to each task or stimulus.[1] Note that due to the preprocessing steps described in Section 5.1, model (14) includes no intercept, nuisance terms, or residual autocorrelation. All priors were specified as described in Section 2, and we used INLA as described in Section 2.2 to obtain posterior estimates of activation for each subject.

We then obtained group-level posterior estimates of activation using the fully Bayes GE model, the joint approach ($L = 50$) and the naive two-level approach described in Section 3. For the motor task study, the fully Bayes GE model was not computationally feasible, so only the joint and two-level approaches were applied. For each method, we used the joint PPM approach described in Section 2.3 to identify group-level areas of activation in response to each task. Three activation thresholds were considered: $\gamma = 0\%$, 0.5% and 1% of the baseline signal.

For comparison, we also applied the classical GLM method to the data. For each subject, we first spatially smoothed the data using the Connectome Workbench geodesic Gaussian surface smoothing algorithm with 6mm FWHM (Glasser et al. 2013). For further comparison, we fit the model with the original, unsmoothed data, as in the Bayesian GLM. Prior to model fitting we performed prewhitening, scaling, centering and nuisance regression as described above. After fitting each subject-level model, we then used the subject-level estimates in an inverse variance weighted least squares model to obtain group-level estimates of activation. Finally, we identified group-level regions of activation by performing a t-test at every vertex, accounting for multiple comparisons with FDR correction and FWER correction as described in the simulation study, except that the permutation test was based on 100 resamplings of each subject's prewhitened time series.

[1] For the model within each hemisphere, the two ipsilateral tasks (e.g., left finger tapping and left toe wiggling for the left hemisphere model) were excluded for computational purposes. Since activation due to lateral tasks is primarily expected in the contralateral hemisphere, these tasks in theory can be ignored within the ipsilateral hemisphere. Appendix Figure C2 displays subject-level activation estimates for the full model and the model excluding ipsilateral tasks. The estimates appear very similar, suggesting that this approach is reasonable.

5.3 Results

Figures 13 and 14 display subject- and group-level activation estimates for the visual cue and tongue tasks of the motor task study. The remaining motor tasks are shown in Appendix C, and subject-level results for the gambling task are shown in Appendix D. For both studies, activation patterns for each task or condition are similar to those observed in previous work (Barch et al. 2013). The Bayesian GLM estimates at the individual and group level are substantially smoother than the classical GLM estimates, even though the data was only smoothed for the classical GLM. As observed in the simulation study, the two-level group approach tends to result in smoother estimates than the joint approach, which likely reflects oversmoothing. In general, the subject-level estimates tend to be noisy, due to the high levels of noise in fMRI data and relatively short duration of the tasks performed for each individual subject (30 seconds for the visual cue and 24 seconds for each motor task).

Figure 15 shows estimated areas of activation for the motor task study using each method. The areas of activation based on $\gamma = 0\%$ in the Bayesian GLM are much larger than those based on FWER correction and even those based on FDR correction. This suggests a gain in power while maintaining strict false positive control, as observed in the simulation study. The different activation thresholds available through the Bayesian GLM provide different interpretations, with $\gamma = 0\%$ including more subtle activations seen in red in Figure 14 and $\gamma = 0.5\%$ or 1% resulting in more conservative areas that correspond to the highly activated areas shown in yellow in Figure 14.

Figure 16 displays group-level estimates and areas of activation for the "loss" condition of the gambling task study using the fully Bayes GE model and the proposed joint approach. The two approaches give very similar results, with the joint approach resulting in slightly smoother estimates and slightly larger areas of activation. The "win" and "neutral" conditions are displayed in Appendix D and show similar overall patterns of activation and give similar results using either method.

These results illustrate the benefits of using the proposed spatial Bayesian framework to account for spatial dependencies in fMRI task activation studies. Compared with the classical GLM, the proposed Bayesian GLM approach results in smoother subject- and group-level estimates of activation and greater power to detect true activations. The proposed joint group

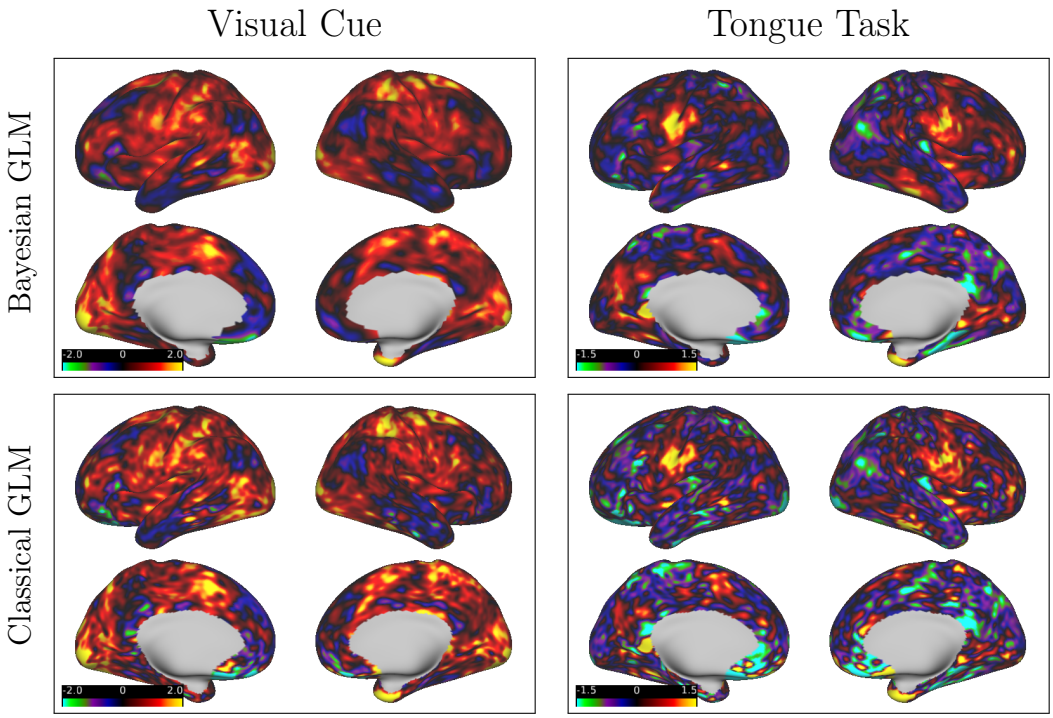

Figure 13: Subject-level estimates of activation amplitude for the visual cue and tongue tasks of the motor study, based on the classical and Bayesian approaches. For the Bayesian GLM, the estimates displayed are posterior means. Classical GLM results are based on smoothed data; results based on unsmoothed data are given in Appendix C.

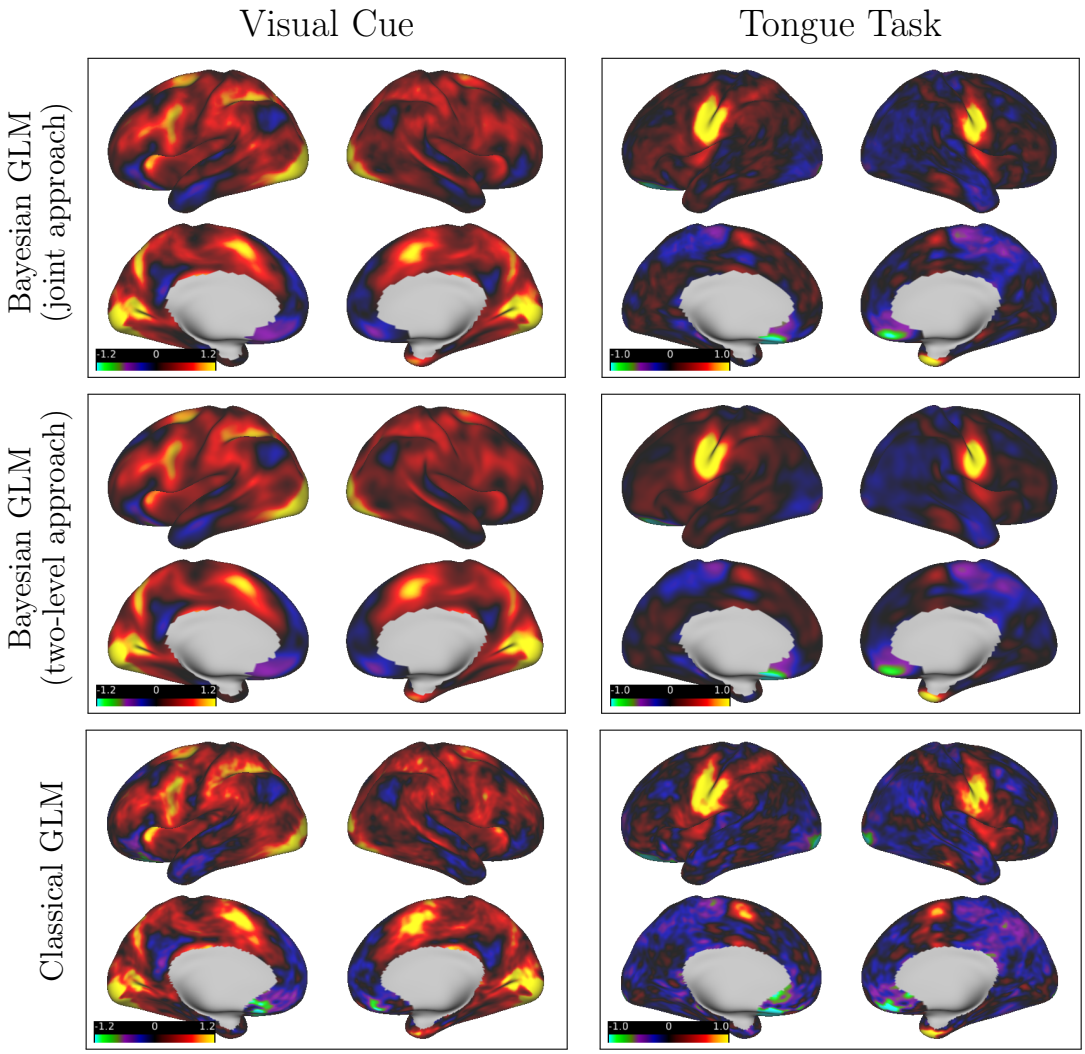

Figure 14: Group-level estimates of activation amplitude for the visual cue and tongue tasks of the motor study, based on the classical and Bayesian approaches. Results for both the joint and two-level Bayesian multi-subject modeling approaches are displayed. As expected, the two-level approach tends to result in oversmoothed activation estimates. Classical GLM results are based on smoothed data; results based on unsmoothed data are given in Appendix C.

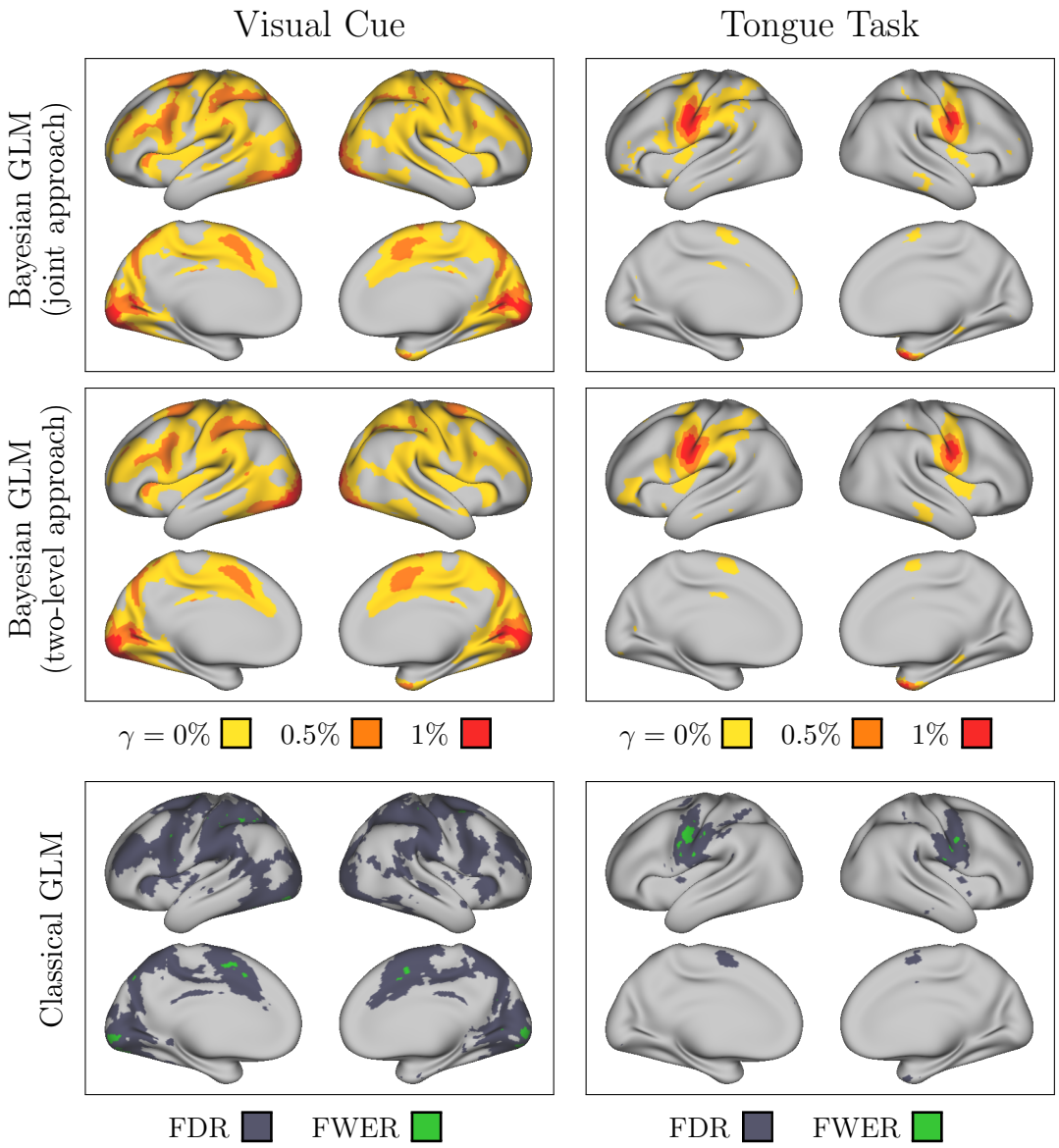

Figure 15: Group-level regions of activation for the visual cue and tongue tasks of the motor study at significance level 0.01, based on the classical and Bayesian approaches. Classical GLM results are based on smoothed data; results based on unsmoothed data are given in Appendix C.

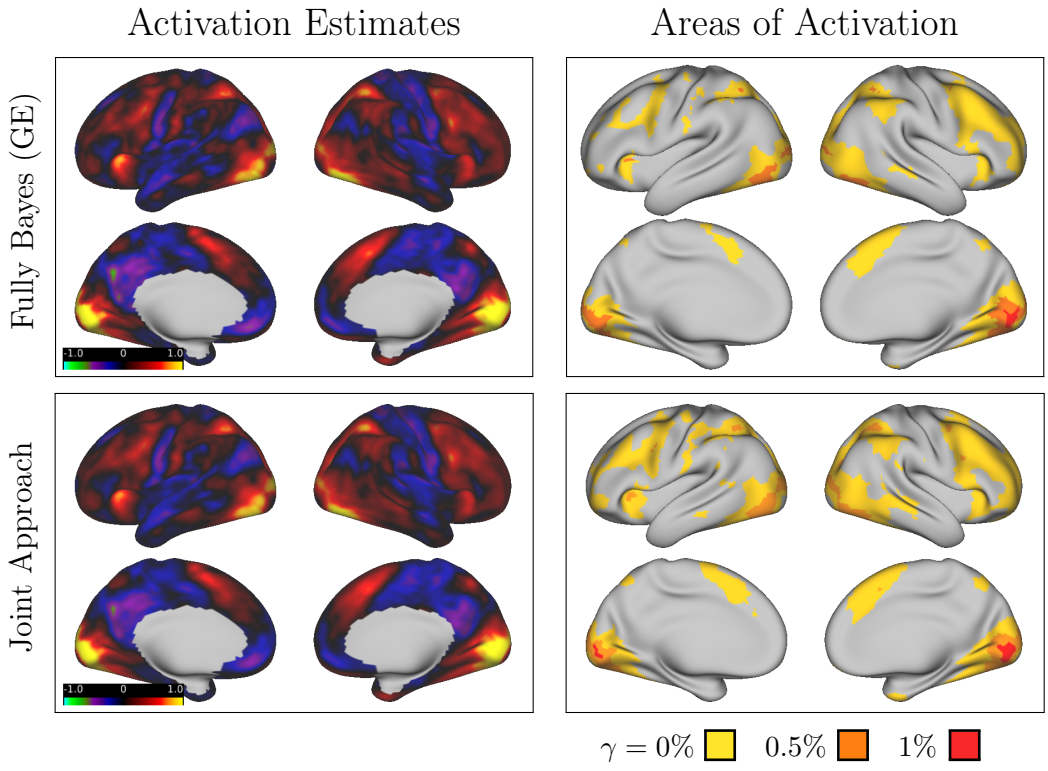

Figure 16: Bayesian group-level estimates and regions of activation for the "loss" condition of the gambling study using the fully Bayes multi-subject GE model and the proposed joint approach. Areas of activation are based on significance level 0.01.

approach gives very similar results to the fully Bayesian group-effects model and provides a computationally efficient and scalable option for group analysis. Our results also illustrate the significant computational advantage of the proposed approach, compared with previously proposed spatial Bayesian models for volumetric fMRI data. Using the proposed methods, whole brain single-subject and group-level spatial Bayesian is quite computationally feasible, requiring less than 1 hour per hemisphere for an individual subject and approximately 12 hours for a group-effects model including 20 subjects in the gambling study. Detailed computation times for the proposed single-subject and multi-subject Bayesian GLM methods, as well as the classical GLM, are reported in Appendix E.

6 DISCUSSION

In this article, we have proposed a novel spatial Bayesian GLM approach for analysis of cortical surface fMRI (cs-fMRI) data, including a joint posterior probability thresholding method for identifying areas of activation and a novel and scalable multi-subject modeling approach. Our three primary contributions—applicability to cs-fMRI, joint posterior probability thresholding, and multi-subject spatial Bayesian modeling—represent major advances for the neuroimaging community, which is interested in obtaining accurate estimates of task activation for individuals and groups of subjects, identifying areas of activation in a way that achieves accurate false positive control and high power, and utilizing cs-fMRI data to provide improved localization and visualization of areas of the brain responsible for various functions. The performance of the proposed Bayesian approach has been validated through simulation studies and motor and gambling task fMRI studies from the Human Connectome Project.

Although the classical GLM is nearly universally practiced, its pitfalls, which include reliance on ad-hoc data smoothing methods, loss of power due to failure to fully leverage spatial dependencies, and the need for multiple comparisons correction techniques that may fail to provide accurate control over false positive rates (Eklund et al. 2012; 2016), are well-known within the neuroimaging community. While spatial Bayesian models have long held promise as a way to address these shortcomings, previously proposed spatial Bayesian models for fMRI analyis have only been applicable to volumetric fMRI data and have not seen wide adoption by the neuroimaging community. This is due in large part to fundamental shortcomings of the data, which

exhibits a complex spatial dependence structure that departs from the parametric assumptions of the models. Here, we have proposed the first spatial Bayesian modeling approach for cs-fMRI data, which is much more compatible with spatial modeling assumptions. We also provide the first use of INLA for Bayesian computation in a spatial Bayesian GLM for fMRI analysis, yielding much more accurate posterior estimates and inference versus variational Bayes techniques.

The proposed methods also provide major computational benefits, with computation time of less than 1 hour per subject and hemisphere in our analyses. By contrast, spatial Bayesian models for volumetric fMRI have been designed to be applied separately to each 2-dimensional slice of the brain volume due to computational limitations, and a recently proposed whole-brain spatial Bayesian model for volumetric fMRI required approximately 100 hours of computation time (Sidén et al. 2017). Our proposed approach is also fast enough to applied to data from several subjects, enabling direct multi-subject analysis in some settings.

These computational gains are the result of several factors, including use of the INLA method and the fast, C-based `R-INLA` implementation with the `PARDISO` parallel sparse matrix library. Further, the dimensionality of cs-fMRI data is lower, with approximately 30,000 vertices per hemisphere compared with over 200,000 voxels within the brain at a 2mm isotropic resolution. The spatial resolution of cs-fMRI can also be coarsened without great loss of information, since data points can be interpolated along the cortical surface, along which neuronal activation tends to vary slowly. Dimension reduction in volumetric fMRI, by contrast, often results in blurring boundaries between tissue classes and, due to cortical folding, mixing signals from distinct areas of the cortex.

Identifying areas of activation in a way that achieves both accurate false positive control and high power is an important and ongoing issue in neuroimaging research. Previously proposed spatial Bayesian models have relied on the marginal posterior distribution at each location to identify regions of activation, given the computational burden associated with using the full joint posterior distribution. Here, we adopt a computationally efficient excursions set method that identifies areas of activation based on the joint posterior distribution for each latent field. This has the result of achieving accurate false positive control without the need for multiple comparisons correction. This may result in increased power, since multiple comparisons correction techniques for fMRI analysis tend to be overly conservative. One limitation of the proposed approach is

that the estimated excursion set may include isolated regions containing only a single voxel or vertex, which are unlikely to be truly activated in isolation. These could be excluded from the excursion set, thereby guaranteeing that the remaining regions have at least the target activation probability, but this would be overly conservative. In the future, we will develop joint posterior thresholding methods that avoid spurious activations while maintaining the target level of false positive control.

Finally, we propose several approaches to multi-subject analysis, a major gap in the existing spatial Bayesian GLM literature. Due to the computational efficiency of the proposed methods, it is possible in some cases to estimate a fully Bayesian group model. However, the total computation time and memory requirements of a fully Bayesian group model would likely be prohibitive in many settings, particularly for large group studies. Therefore, we also propose a novel joint modeling approach, which combines the results of subject-level models in a principled way to provide posterior inference on the group-level quantities of interest. Our simulated data analysis shows this approach to have good accuracy compared with the fully Bayesian approach, and the two approaches give very similar results in our experimental data analysis. This represents a major advance for fMRI task analysis, since it provides the first computationally feasible technique for multi-subject Bayesian spatial modeling and is scalable through parallelization.

Several limitations of the proposed methods should be noted. First, while `R-INLA` provides a convenient, sophisticated and in some ways highly efficient implementation of the INLA method, it is a general tool that may not be optimal for our particular context. In particular, it is not optimized for high-dimensional data, which increases the computational burden associated with model estimation. A custom software implementation of the methods utilized here may be able to achieve greater computational efficiency. This would make it possible to analyze more high-dimensional data, allowing us to perform a lesser degree of data resampling and include more subjects in a fully Bayesian group model. A customized implementation would also allow for targeted methodological modifications to the INLA method for greater accuracy, flexibility and computational efficiency. Future work will therefore focus on developing a custom software implementation optimized for spatial Bayesian modeling of fMRI data.

An alternative approach for the Bayesian computation would be to employ Markov chain Monte Carlo (MCMC). While INLA tends to be faster than MCMC in general since no sampling

is performed, it may be possible to develop a customized MCMC sampler that takes advantage of the structure of the problem in a way that the R-INLA implementation of INLA, as a general tool, is not capable of. The use of MCMC for Bayesian computation in spatial Bayesian models for cs-fMRI is an important area of future research.

Second, the SPDE priors we utilize are suitable for modeling stationary and isotropic spatial processes. In volumetric fMRI data, it is well-established that amplitude fields tend to present non-stationary features such as varying degrees of activation (e.g., Harrison et al. 2008, Yue et al. 2010). While we largely avoid these issues through the use of cortical surface fMRI data, some non-stationary features may persist. For example, activated areas tend to be less smooth than background areas. Lindgren et al. (2011) introduced non-stationarity to the SPDE models by allowing κ and τ in (2) to depend on location, and Bolin and Lindgren (2011) and Fuglstad et al. (2014) generalized further using nested directional operators and non-isotropic Laplacians, respectively. Future work will investigate the feasibility and performance of non-stationary SPDE models in the Bayesian GLM framework.

Third, several simplifications to the model were made in order to improve computational efficiency. These include prewhitening to remove temporal autocorrelation in the residuals and elimination of nuisance signals and the baseline field, since these are not of primary scientific interest. Both of these steps are analogous to an empirical Bayes approach. While more formal approaches would be possible in theory, we believe that these steps represent reasonable compromises with significant computational benefits. The handling of residual autocorrelations in task fMRI is a highly controversial and current topic within the neuroimaging community. This is in part due to the growing prevalence of fast fMRI acquisitions, which tend to exacerbate residual autocorrelations. Further, recent research and our own observations have shown that residual autocorrelation tends to vary spatially, often quite dramatically. In theory, it would be possible to model such residual autocorrelation patterns as smoothly-varying latent fields within a spatial Bayesian model. While this would be computationally prohibitive within the current framework, it represents an interesting and relevant topic for future research.

Finally, while the proposed methods assume that spatial dependence is inversely related to distance, a more complex network structure involving long-range connections may exist in the data. For example, during motor tasks certain cerebellar regions tend to co-activate along with

cortical motor regions, since the cerebellum plays an important role in motor control. Future work should aim to incorporate such long-range dependencies between different brain regions during task activity. We will also investigate the potential of an empirical Bayesian framework employing estimates of long-range and short-range dependence to reduce the computational burden compared with the fully Bayesian approach proposed here. Additionally, the distances utilized in this paper are radial distances along the spherical surface. While the spherical representation of the cortical surface of each hemisphere provides a simple geometric setting while minimizing distance distortions, some level of distortion is inevitable. In the future, we will develop techniques to utilize geodesic distances along the true cortical surface of each subject to more accurately capture distance-related spatial dependencies.

ACKNOWLEDGEMENTS

Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. This research was supported in part by NIH grants P41 EB015909, R01 EB016061 and R01 EB027119 from the National Institute of Biomedical Imaging and Bioengineering, PSC-CUNY Research Award 67192-00 45, and Swedish Research Council grant 2016-04187.

SUPPLEMENTARY MATERIALS

Appendix: File containing proofs and additional figures. (PDF)

Code: Set of bash, MATLAB and R scripts used to perform the analysis. (zip file)

REFERENCES

Adler, R. J. (1981). *The Geometry of Random Fields*. Wiley, New York.

Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., Glasser, M. F., Curtiss, S., Dixit, S., Feldt, C. et al. (2013). Function in the human connectome: task-fmri and individual differences in behavior. *Neuroimage* **80**, 169–189.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 289–300.

Bishop, C. M. (2006). Pattern recognition. *Machine Learning* **128**, 1–58.

Bolin, D. and Lindgren, F. (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *The Annals of Applied Statistics* **5**, 523–550.

Bolin, D. and Lindgren, F. (2013). A comparison between Markov approximations and other methods for large spatial data sets. *Computational Statistics and Data Analysis* **61**, 7–21.

Bolin, D. and Lindgren, F. (2015). Excursion and contour uncertainty regions for latent Gaussian models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **77**, 85–106.

Bolin, D. and Lindgren, F. (2018). Calculating probabilistic excursion sets and related quantities using excursions. *Journal of Statistical Software* **86**.

Dale, A. M., Fischl, B. and Sereno, M. I. (1999). Cortical surface-based analysis I: Segmentation and surface reconstruction. *NeuroImage* **9**, 179–194.

Eklund, A., Andersson, M., Josephson, C., Johannesson, M. and Knutsson, H. (2012). Does parametric fMRI analysis with SPM yield valid results? An empirical study of 1484 rest datasets. *NeuroImage* **61**, 565–578.

Eklund, A., Nichols, T. E. and Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences* 201602413.

Fischl, B. (2012). FreeSurfer. *NeuroImage* **62**, 774–781.

Fischl, B., Sereno, M. I. and Dale, A. M. (1999). Cortical surface-based analysis II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage* **9**, 195–207.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A. and Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fmri): use of a cluster-size threshold. *Magnetic Resonance in medicine* **33**, 636–647.

Friston, K. and Penny, W. (2003). Posterior probability maps and SPMs. *NeuroImage* **19**, 1240–1249.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. and Turner, R. (1998). Event-related fMRI: characterizing differential responses. *NeuroImage* **7**, 30–40.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D. and Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping* **2**, 189–210.

Fuglstad, G.-A. . A., Lindgren, F., Simpson, D. and Rue, H. (2014). Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy. *Statistica Sinica* (accepted).

Genovese, C. R., Lazar, N. A. and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* **15**, 870–878.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R. et al. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124.

Guttorp, P. and Gneiting, T. (2006). Studies in the history of probability and statistics XLIX On the Matérn correlation family. *Biometrika* **93**, 989–995.

Harrison, L. M., Penny, W., Daunizeau, J. and Friston, K. J. (2008). Diffusion-based spatial priors for functional magnetic resonance images. *NeuroImage* **41**, 408–423.

HCP (2016). Task fMRI files and protocol details. [Online; accessed 08-November-2016].

Ishwaran, H. and Rao, J. S. (2003). Detecting differentially expressed genes in microarrays using Bayesian model selection. *Journal of the American Statistical Association* **98**, 438–455.

Lindgren, F. and Rue, H. (2015). Bayesian spatial modelling with R-INLA. *Journal of Statistical Software* **63**.

Lindgren, F., Rue, H. and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**, 423–498.

Lindquist, M. A. (2008). The statistical analysis of fMRI data. *Statistical Science* **23**, 439–464.

Lindquist, M. A. and Mejia, A. (2015). Zen and the art of multiple comparisons. *Psychosomatic Medicine* **77**, 114.

Lindquist, M. A., Spicer, J., Asllani, I. and Wager, T. D. (2012). Estimating and testing variance components in a multi-level GLM. *NeuroImage* **59**, 490–501.

Marchini, J. and Presanis, A. (2004). Comparing methods of analyzing fMRI statistical parametric maps. *NeuroImage* **22**, 1203–1213.

Marcus, D., Harwell, J., Olsen, T., Hodge, M., Glasser, M., Prior, F., Jenkinson, M., Laumann, T., Curtiss, S. and Van Essen, D. (2011). Informatics and data mining tools and strategies for the Human Connectome Project. *Frontiers in Neuroinformatics* **5**, 4.

Martins, T. G., Simpson, D., Lindgren, F. and Rue, H. (2013). Bayesian computing with INLA: New features. *Computational Statistics and Data Analysis* **67**, 68 – 83.

Nichols, T. and Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical Methods in Medical Research* **12**, 419–446.

Nichols, T. E. and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping* **15**, 1–25.

Penny, W. D., Trujillo-Barreto, N. J. and Friston, K. J. (2005). Bayesian fMRI time series analysis with spatial priors. *NeuroImage* **24**, 350–362.

Poldrack, R. A., Mumford, J. A. and Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. Cambridge University Press.

Poline, J.-B. and Mazoyer, B. M. (1993). Analysis of individual positron emission tomography activation maps by detection of high signal-to-noise-ratio pixel clusters. *Journal of Cerebral Blood Flow & Metabolism* **13**, 425–437.

Rue, H., Martino, S. and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B: Statistical Methodology* **71**, 319–392.

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P. and Lindgren, F. K. (2016). Bayesian computing with INLA: A review. *arXiv preprint arXiv:1604.00860* .

Sidén, P., Eklund, A., Bolin, D. and Villani, M. (2017). Fast Bayesian whole-brain fMRI analysis with spatial 3D priors. *NeuroImage* **146**, 211–225.

Smith, S. M. and Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* **44**, 83–98.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., WU-Minn HCP Consortium et al. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage* **80**, 62–79.

Wager, T. D., Hernandez, L. and Lindquist, M. A. (2009). Essentials of functional neuroimaging. *Handbook of neuroscience for the behavioral sciences* .

Wang, B. and Titterington, D. (2005). Inadequacy of interval estimates corresponding to variational Bayesian approximations. In *Proceedings of the 10th International Workshop on Artificial Intelligence and Statistics* (R. Cowell and Z. Ghahramani, eds.), 373–380, Society for Artificial Intelligence and Statistics.

Woo, C.-W., Krishnan, A. and Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *NeuroImage* **91**, 412–419.

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M. and Smith, S. M. (2004). Multilevel linear modelling for fMRI group analysis using Bayesian inference. *NeuroImage* **21**, 1732–47.

Worsley, K. J. and Friston, K. J. (1995). Analysis of fMRI time-series revisited–again. *NeuroImage* **2**, 173–181.

Worsley, K. J., Liao, C., Aston, J., Petre, V., Duncan, G., Morales, F. and Evans, A. (2002). A general statistical analysis for fMRI data. *NeuroImage* **15**, 1–15.

Yue, Y., Loh, J. M. and Lindquist, M. A. (2010). Adaptive spatial smoothing of fMRI images. *Statistics and Its Interface* **3**, 3–13.

A Bayesian General Linear Modeling Approach to Cortical Surface fMRI Data Analysis: **Appendix**

Amanda F. Mejia[a1], Yu Ryan Yue[b1], David Bolin[c], Finn Lindgren[d] and Martin A. Lindquist[e]

[a]*Indiana University, Bloomington, IN 47405*
[b]*Baruch College, The City University of New York, New York, NY 10010*
[c]*University of Gothenburg, Gothenburg, Sweden*
[d]*The University of Edinburgh, Edinburgh, UK*
[e]*Johns Hopkins University, Baltimore, MD 21205*

[1]A. Mejia and Y. Yue contributed equally to this work.

Appendix A: Proof of Equation (10)

$$
\begin{aligned}
\pi(\boldsymbol{y} \mid \boldsymbol{\theta}) &= \int \pi(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m \mid \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m, \boldsymbol{\theta}) \pi(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m \mid \boldsymbol{\theta}) d\boldsymbol{\beta}_1 \cdots \boldsymbol{\beta}_m \\
&= \int \prod_{m=1}^{M} \pi(\boldsymbol{y}_m \mid \boldsymbol{\beta}_m, \boldsymbol{\theta}) \pi(\boldsymbol{\beta}_m \mid \boldsymbol{\theta}) d\boldsymbol{\beta}_1 \cdots \boldsymbol{\beta}_m \\
&= \prod_{m=1}^{M} \int \pi(\boldsymbol{y}_m \mid \boldsymbol{\beta}_m, \boldsymbol{\theta}) \pi(\boldsymbol{\beta}_m \mid \boldsymbol{\theta}) d\boldsymbol{\beta}_m \\
&= \prod_{m=1}^{M} \pi(\boldsymbol{y}_m \mid \boldsymbol{\theta}) \\
&\propto \prod_{m=1}^{M} \pi(\boldsymbol{\theta} \mid \boldsymbol{y}_m) \pi(\boldsymbol{\theta})^{-1} \\
&= \pi(\boldsymbol{\theta})^{-M} \prod_{m=1}^{M} \pi(\boldsymbol{\theta} \mid \boldsymbol{y}_m)
\end{aligned}
$$

Appendix B: Simulation Figures

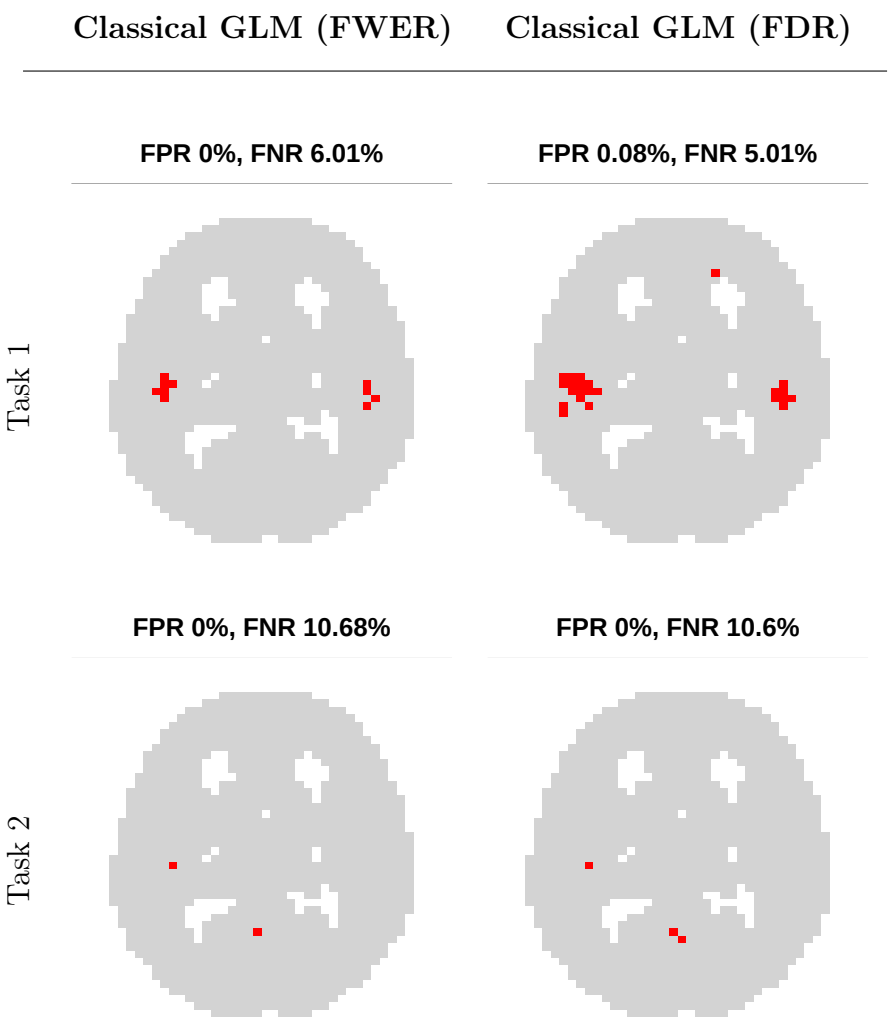

Figure B1: Estimated regions of activation in the simulation study using the classical GLM, based on unsmoothed data. The false positive rate (FPR) and false negative rate (FNR) are reported for each method and activation. Regions of activation are estimated by performing a hypothesis test on the task coefficient at each location, correcting for multiple comparisons through FDR control ($q = 0.01$) and FWER control ($\alpha = 0.01$).

Appendix C: Additional Results, Motor Task Study

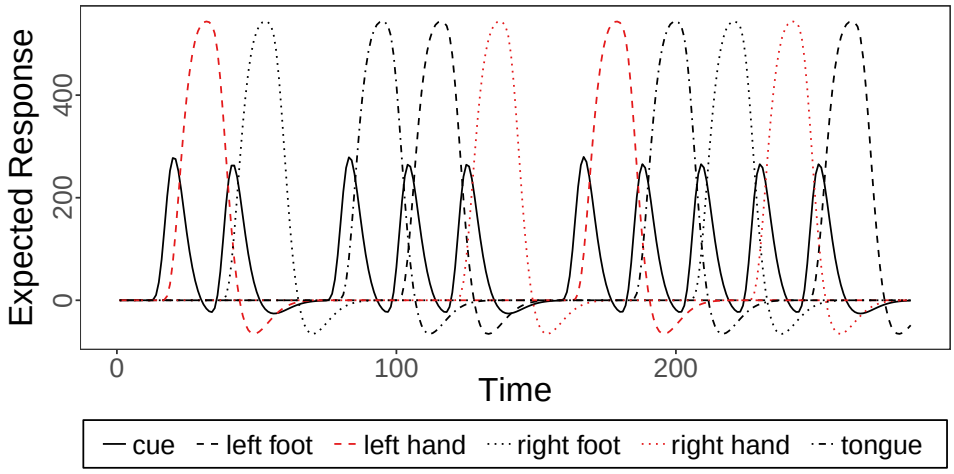

Figure C1: Activation profiles for the visual cue and five motor tasks in the motor study.

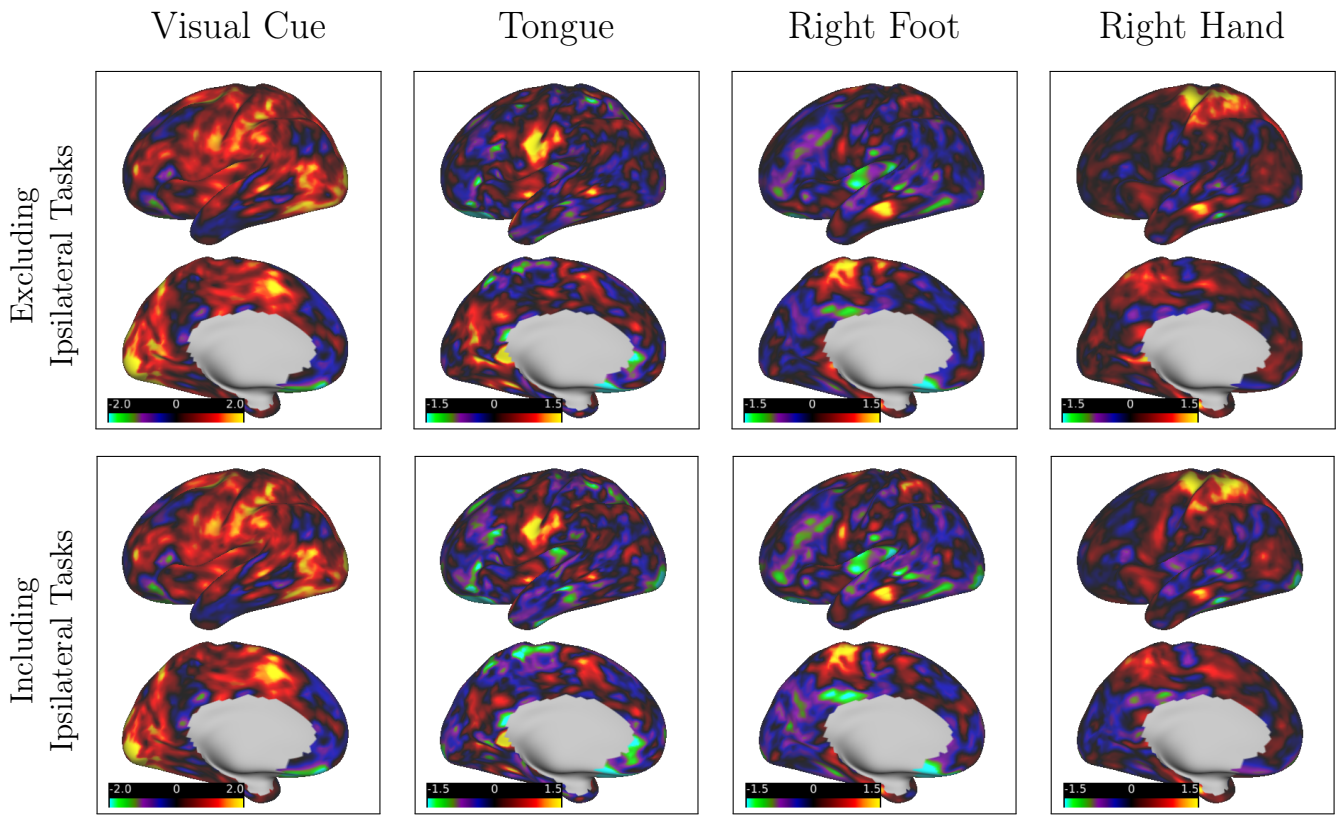

Figure C2: Subject-level estimates of activation amplitude for the left hemisphere Bayesian model, with and without ipsilateral tasks (left finger tapping, left toe wiggling) included in model. For the results of the motor study presented in the paper, we excluded ipsilateral tasks for the model within each hemisphere since activation due to lateral tasks is primarily expected on the contralateral side of the brain. Excluding these two tasks dramatically reduces the computational burden of the model for each hemisphere (see Computation Times section in main text) and, as shown here, produces very similar activation estimates compared with the "full" model.

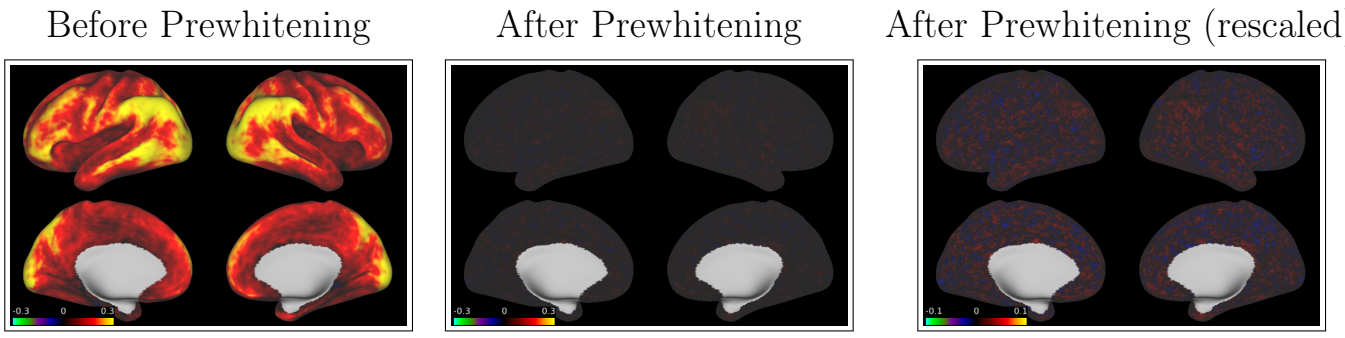

Figure C3: First AR coefficient of the residuals of a classical GLM, before and after prewhitening. Even without multiple iterations, the temporal autocorrelation of the residuals appears to be negligible after prewhitening.

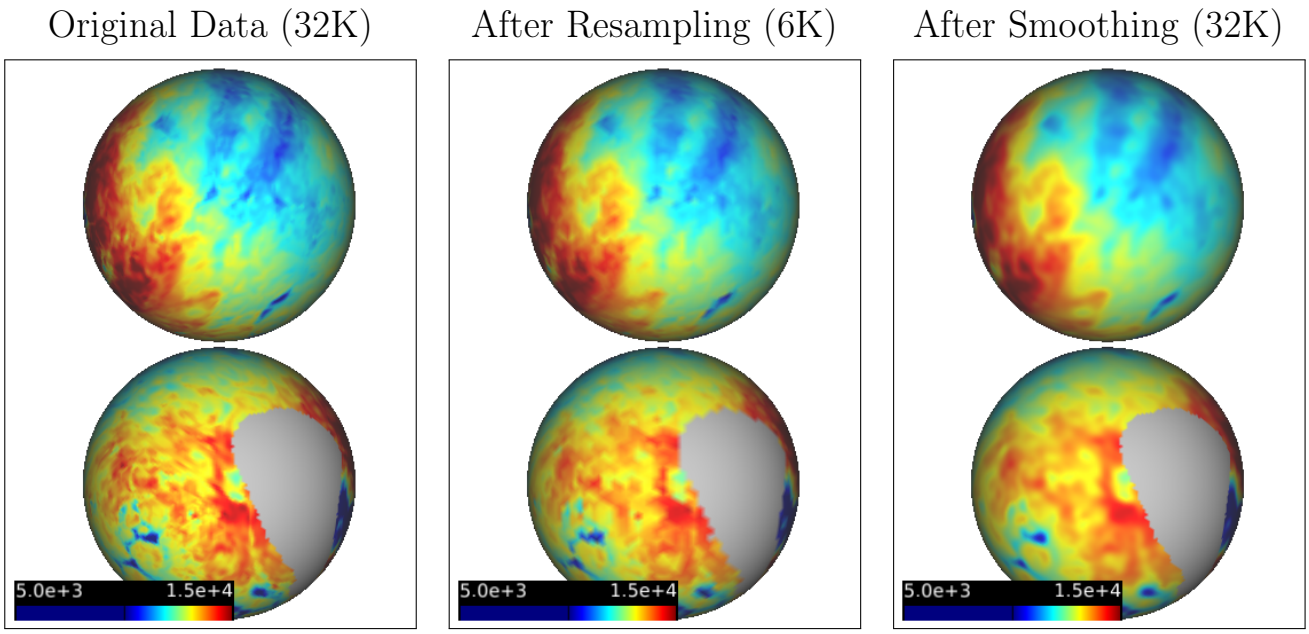

Figure C4: For a single volume of one subject, fMRI BOLD values at the original 32K resolution, after resampling to 6K, and after smoothing the original data with a 6mm FWHM Gaussian kernel. The values are displayed on the spherical surface of the left hemisphere. The data is substantially more blurred due to smoothing versus resampling.

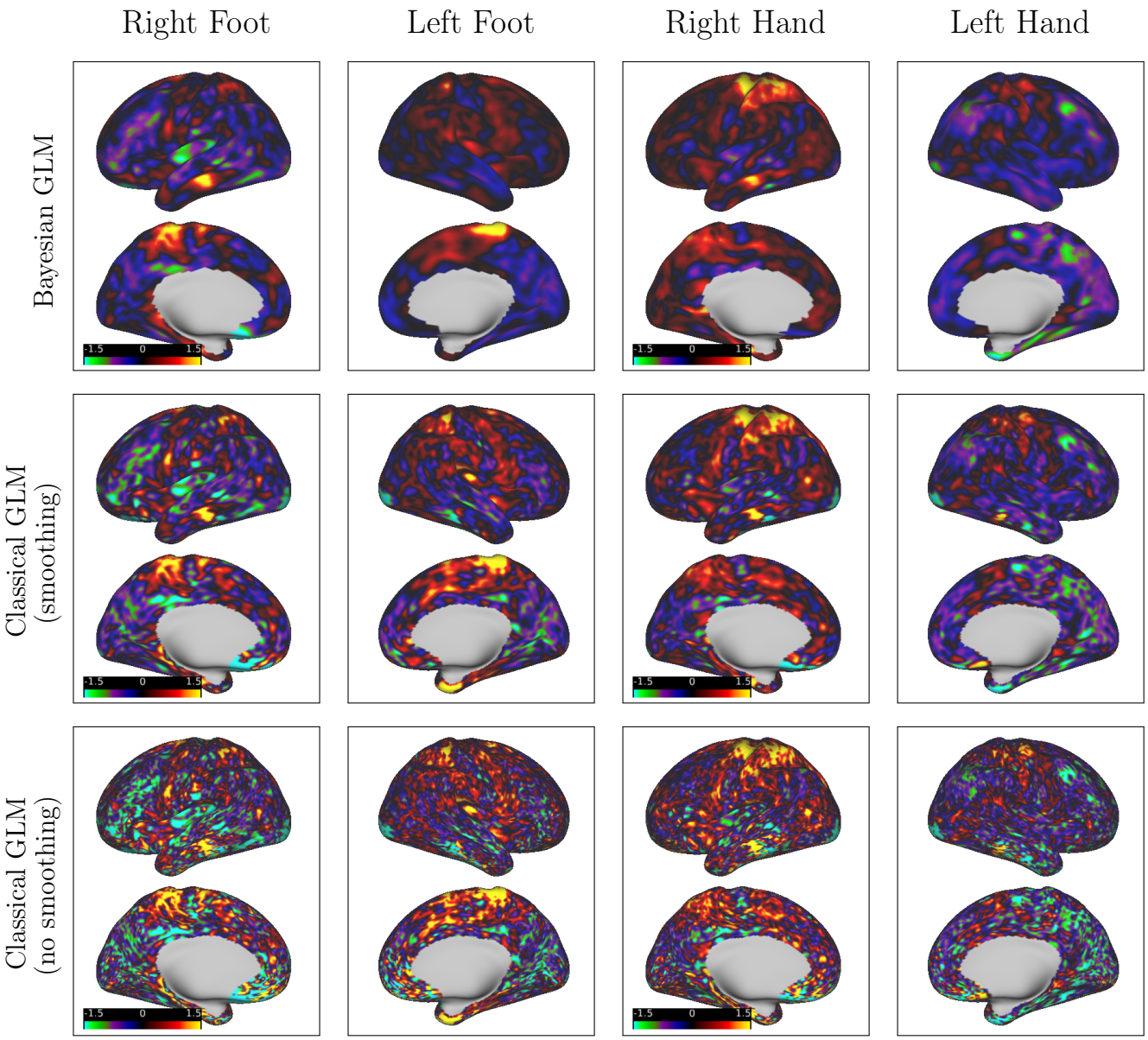

Figure C5: For one randomly selected subject, estimates of activation amplitude for the right foot, left foot, right hand and left hand tasks of the motor study, based on the classical and Bayesian approaches.

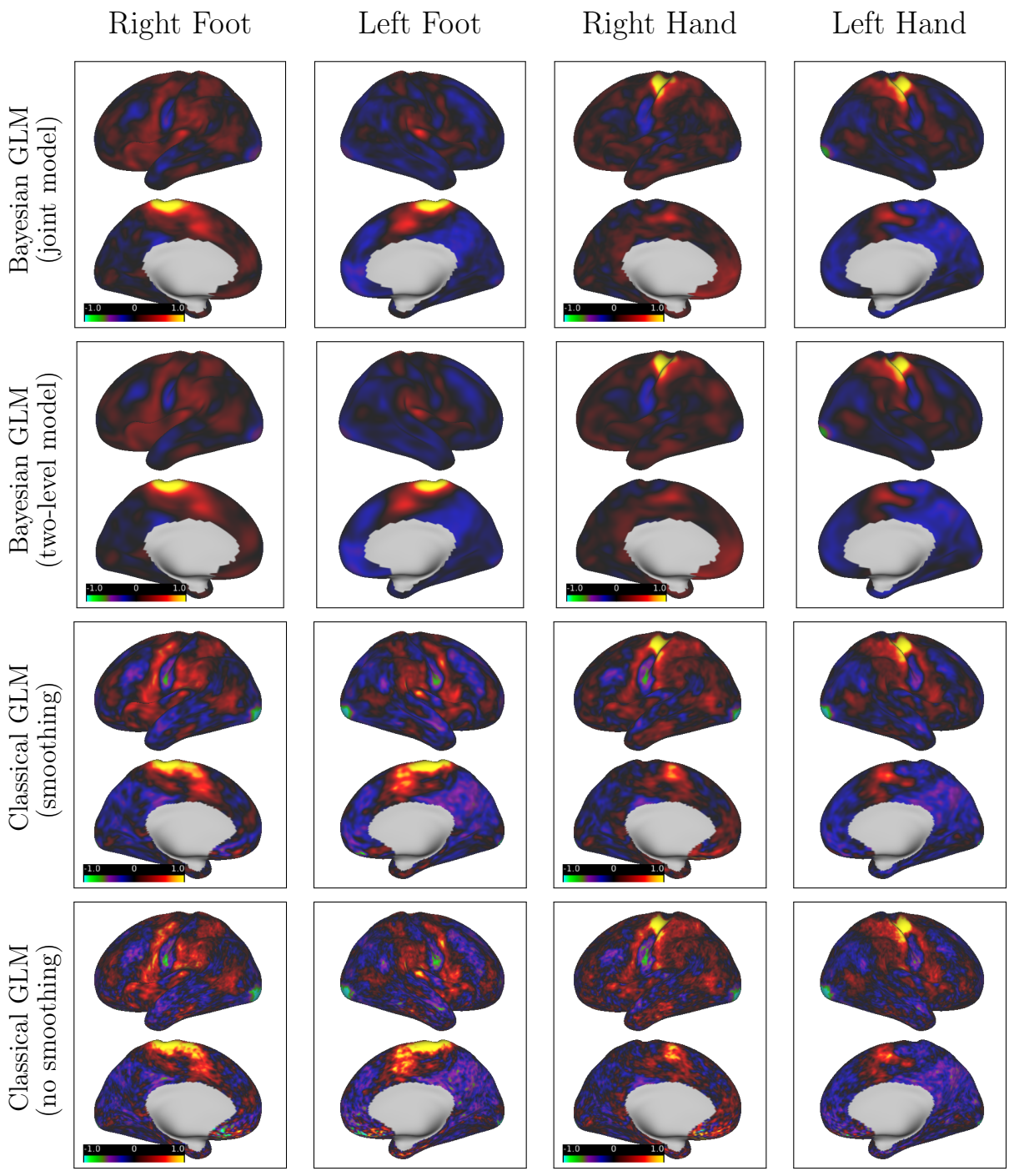

Figure C6: Group-level estimates of activation amplitude for the right foot, left foot, right hand and left hand tasks of the motor study, based on the classical and Bayesian approaches. Results for both the joint and two-level Bayesian multi-subject modeling approaches are displayed. As expected, the two-level approach tends to result in oversmoothed activation estimates.

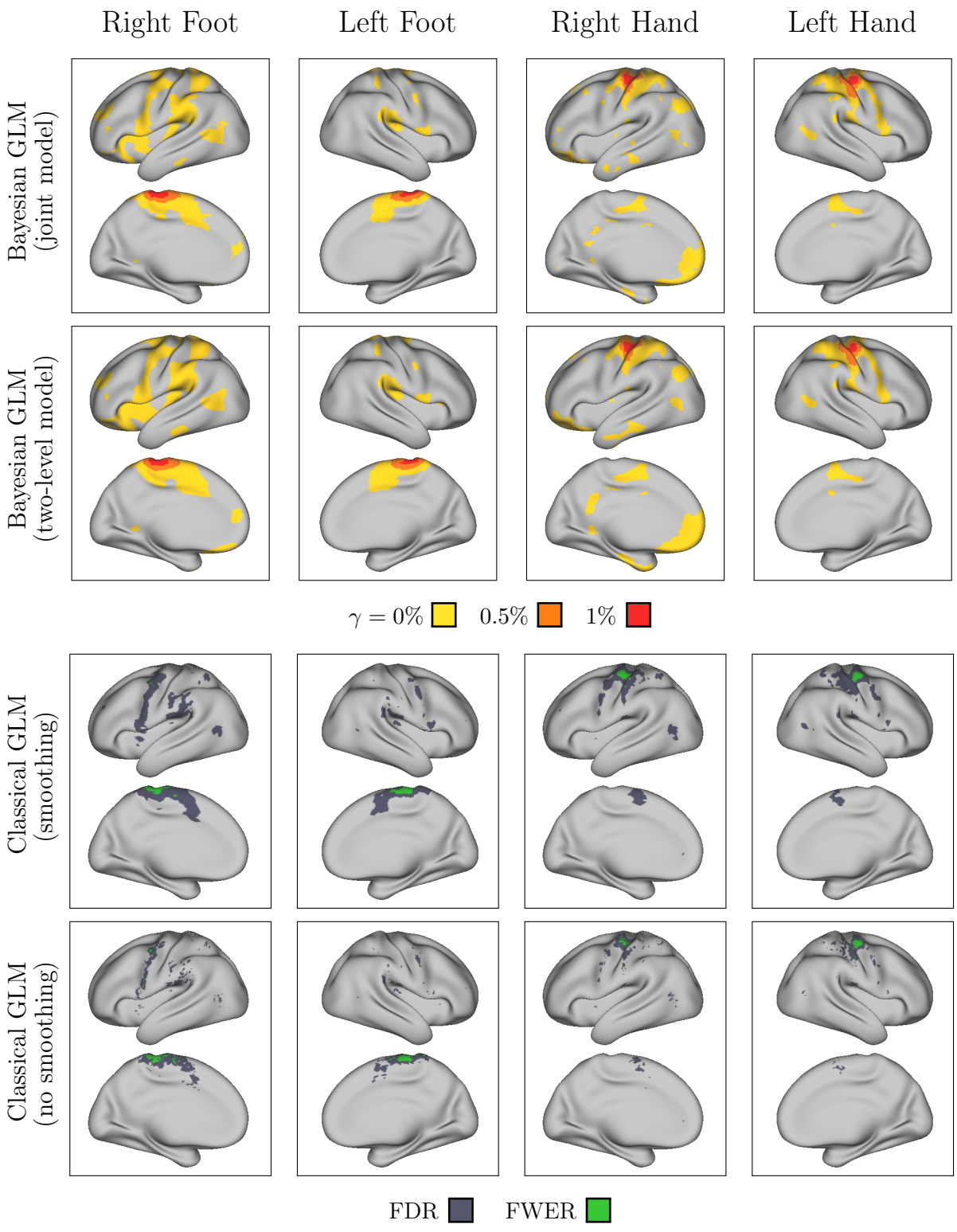

Figure C7: Group-level regions of activation for the right foot, left foot, right hand and left hand tasks of the motor study at significance level 0.01, based on the classical and Bayesian approaches.

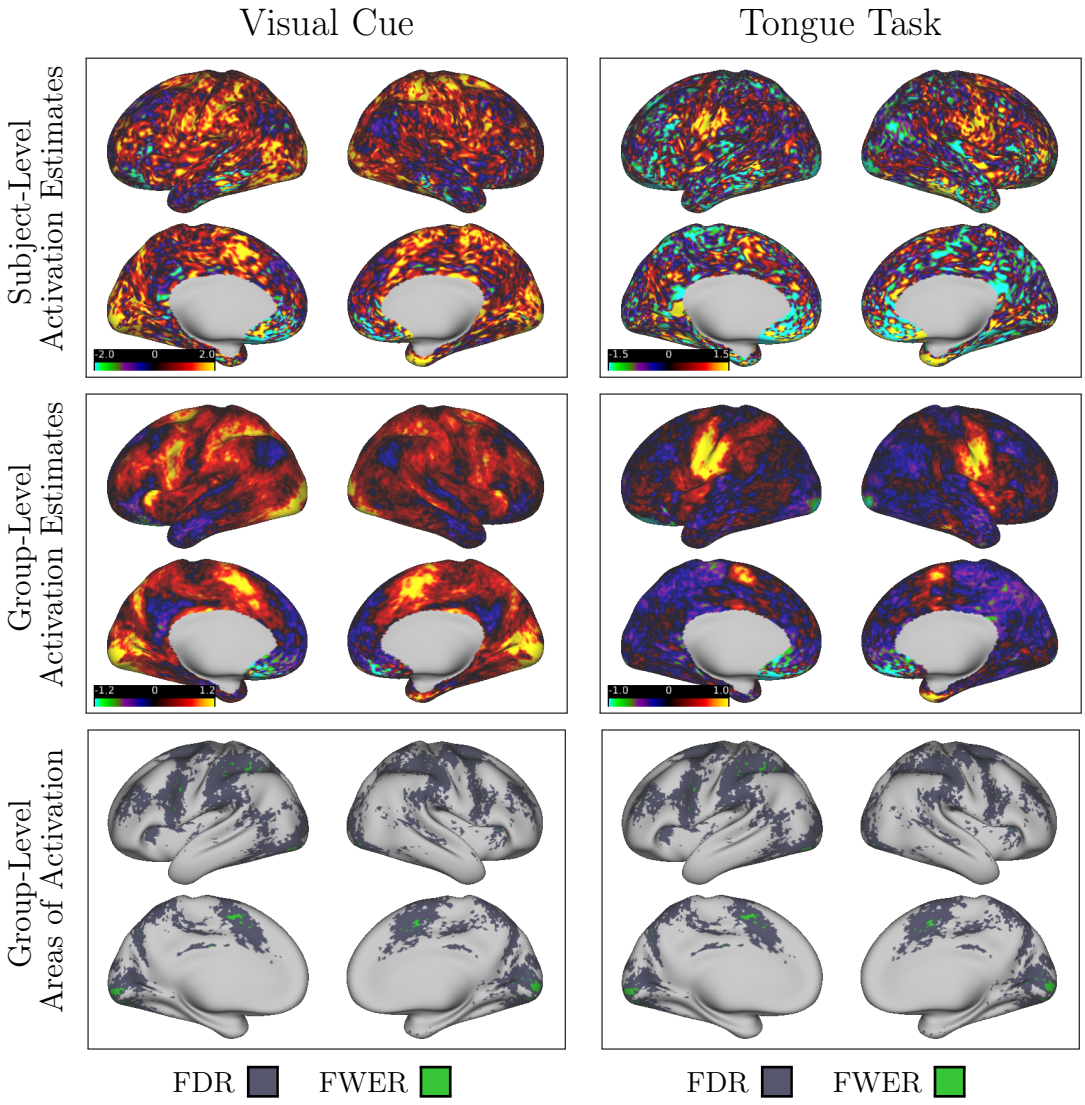

Figure C8: Subject-level estimates of activation amplitude for the visual cue and tongue tasks of the motor study, based on the classical GLM using unsmoothed data.

Appendix D: Results, Gambling Task

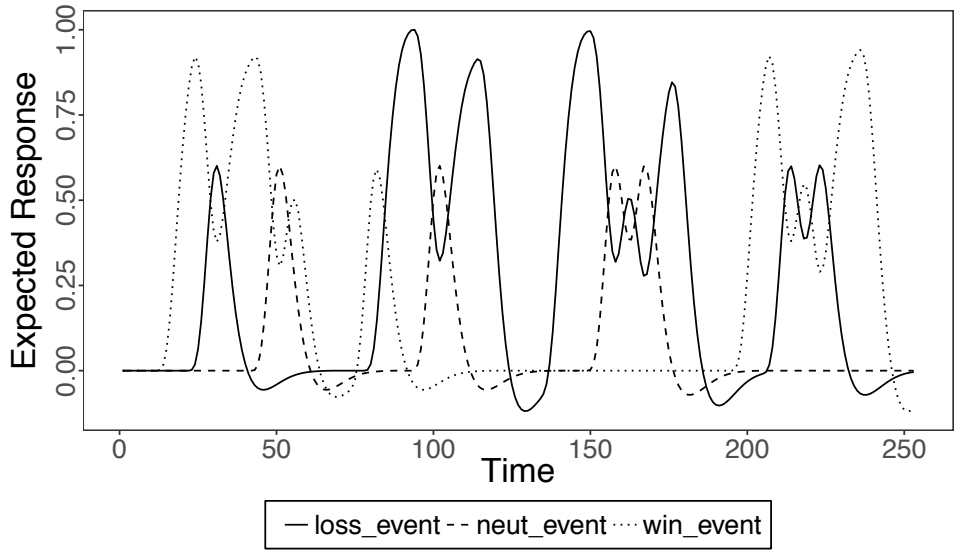

Figure D1: Activation profiles for the gambling task study.

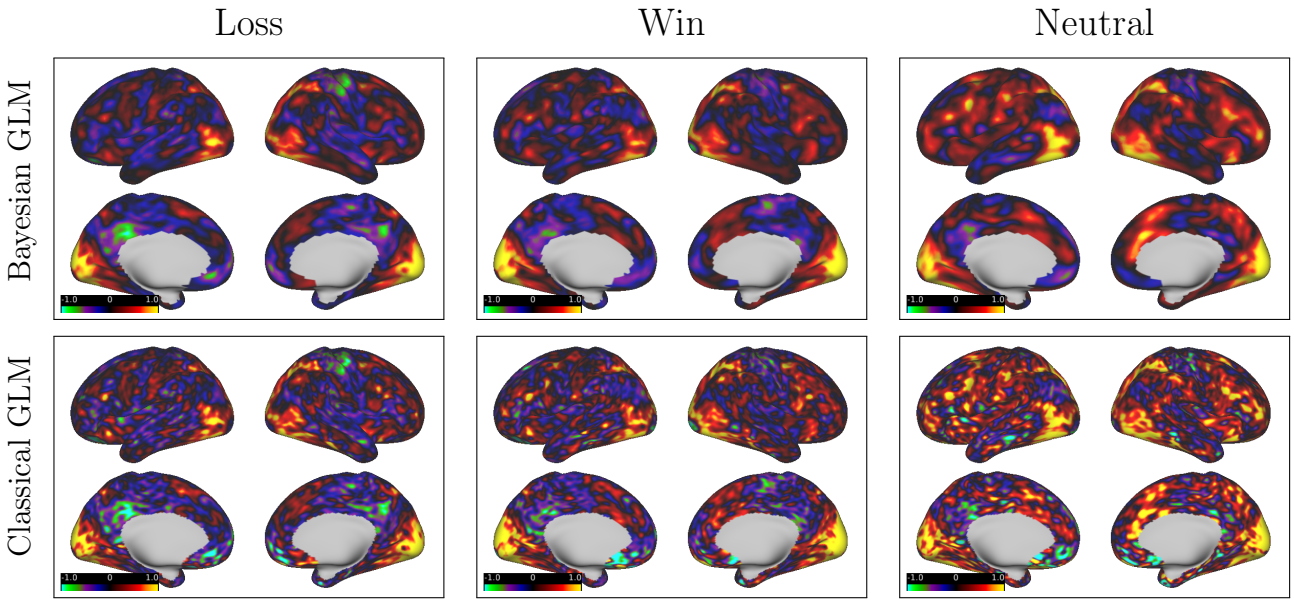

Figure D2: Subject-level estimates of activation amplitude for loss, win and neutral trials in the gambling study using the classical and Bayesian approaches. Classical GLM results are based on smoothed data.

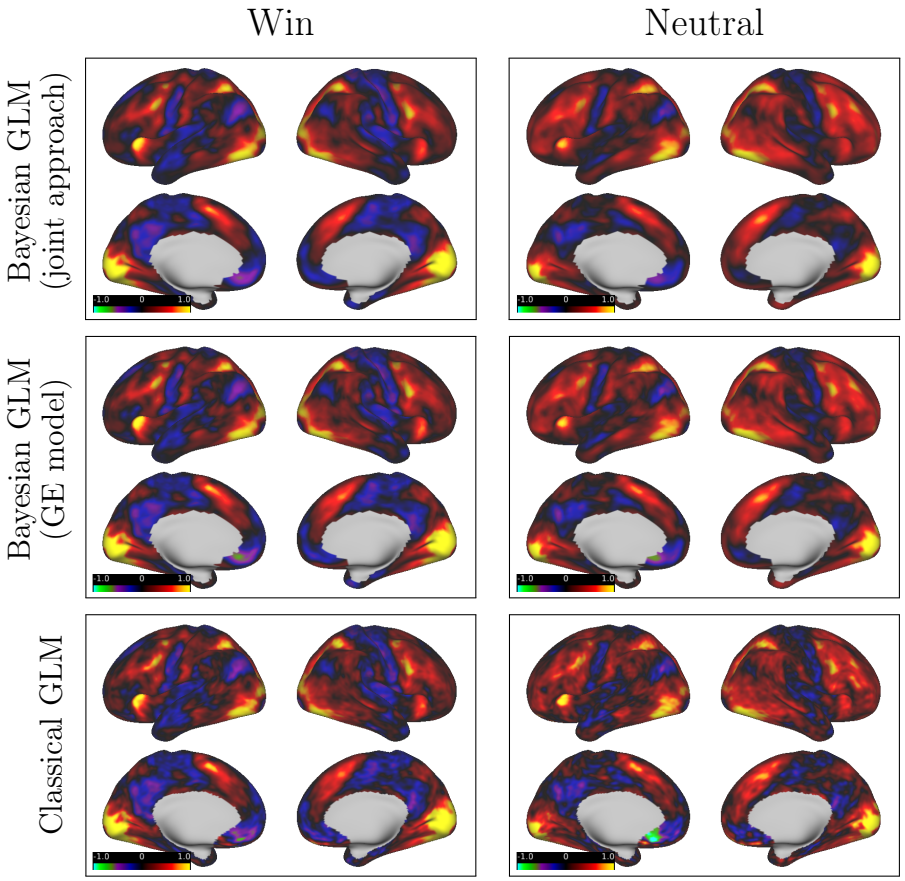

Figure D3: Group-level estimates of activation amplitude for win and neutral trials in the gambling study using the classical and joint Bayesian approaches. All three trial types activate expected areas, including the visual cortex and the insula, which is known to be involved in emotion processing. The win trial has somewhat higher estimates of activation in both areas. Classical GLM results are based on smoothed data.

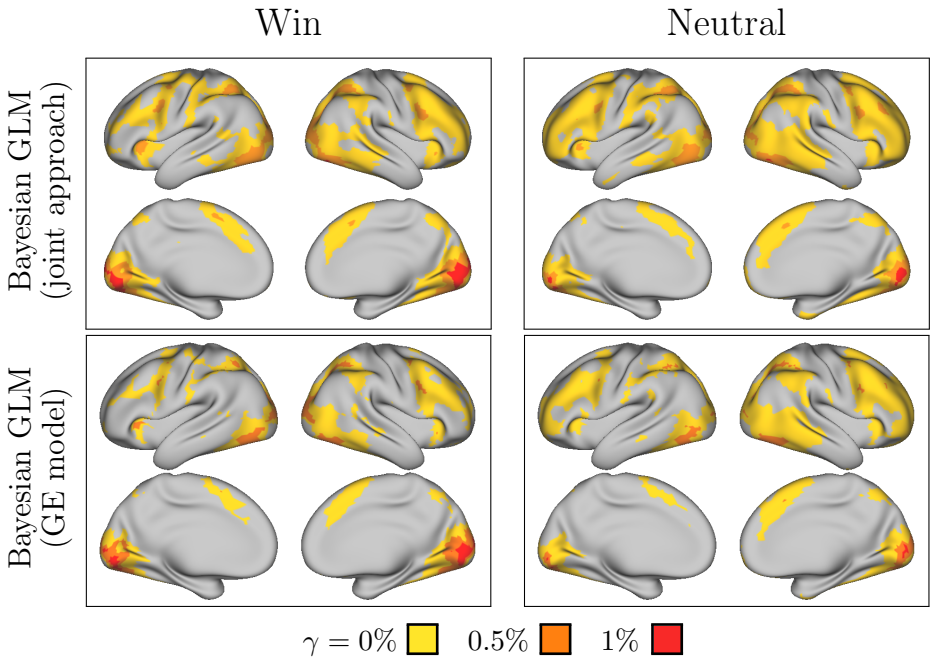

Figure D4: Group-level areas of activation for win and neutral trials in the gambling study using the joint approach and fully Bayesian GE model. Significance level is $\alpha = 0.01$.

Appendix E: Computation Times

The single-subject Bayesian GLM for the motor task study, which included 4 tasks and hence 9 hyperparameters in the model for each hemisphere (two hyperparameters per task, plus one hyperparameter for the residual precision assuming independent errors), required approximately 45 minutes of computation time and 50 gigabytes (GB) of memory per subject and hemisphere. The model including all 6 tasks, resulting in 13 hyperparameters in the model for each hemisphere, required 2 hours of computation time per hemisphere. The gambling task study, which included 3 tasks and therefore 7 hyperparameters, required 30 minutes and 40 GB of memory on average per subject and hemisphere. The difference in computation times for these three models demonstrates the dramatic impact of additional hyperparameters on the computational burden of INLA. Computation time for all models was greatly improved by using the `PARDISO` parallel matrix sparse matrix library. Choosing good starting values for the hyperparameters[1] also substantially improved computation time.

The proposed joint approach for multi-subject analysis, parallelized across samples $\ell = 1, \ldots, L = 50$, required approximately 30 minutes of computation time for the motor task study, including identification of areas of activation for three activation thresholds. Without parallelization, the computation time would have been approximately 6 hours. For the gambling task study, the total computation time using the joint approach was similar at 25 minutes. The two-level group Bayesian approach was more time-consuming due to the need to fit the model for each task in each hemisphere using INLA. Each task required an average of approximately 5 minutes for model fitting and 4-6 minutes for identifying areas of activation at each activation threshold, for a total of approximately 20 minutes per task per hemisphere. The fully Bayesian multi-subject group effect (GE) model for the gambling task study with all 20 subjects required approximately 11 hours of computation time and 270 GB of memory per hemisphere for model fitting, followed by an additional 1 hour for excursion set estimation for each task and activation threshold. The growth in computation time and memory requirements per additional subject was approximately linear for the GE model.

Finally, the single-subject classical GLM took 3h 39m of computation time and required 1.39 GB of memory on average for the motor task study. For the gambling task study, the computation time was 2h 40m on average per subject with similar memory requirements. The relatively high computation time was due to the vertex-wise prewhitening approach, which results in a unique design matrix at each vertex and thus requires looping over all vertices. Refitting the model at each location for each subject using 100 permutations of the prewhitened time points required approximately 1 hour of additional computation time per subject and 1.1 GB of storage space. Performing FWER correction required an additional 3.5 hours of computation time at the group level. Performing FDR correction, by contrast, took less than 1 second at the group level and no additional computation time at the subject level.

[1] We chose -2 for $\log(\tau)$ and 2 for $\log(\kappa)$.